IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Amy S. Lee
Serial No. : 09/606,804
Filed : June 28, 2000
Title : STRESS-RESPONSIVE INDUCTION OF A THERAPEUTIC AGENT AND METHODS OF USE

Art Unit : 1635
Examiner : Brian A. Whiteman

Commissioner for Patents
Washington, D.C. 20231

DECLARATION UNDER 37 C.F.R. § 1.132

I, Amy S. Lee, hereby declare as follows:

1) I am the Amy S. Lee who is named as the sole inventor on the above-identified patent application. I believe I am an original and first inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled "STRESS-RESPONSIVE INDUCTION OF A THERAPEUTIC AGENT AND METHODS OF USE."

2) I am the senior author in a publication with Gadi Gazit, Gene Hung, Xinke Chen and W. French Anderson entitled "USE OF THE GLUCOSE STARVATION-INDUCIBLE GLUCOSE REGULATED PROTEIN 78 PROMOTER IN SUICIDE GENE THERAPY OF MURINE FIBROSARCOMA," in *Cancer Research*, **59**:3100-3106 (July 1, 1999).

3) The invention described and claimed in the subject application was conceived by myself and reduced to practice under my direction and supervision.

CERTIFICATE OF MAILING BY FIRST CLASS MAIL

I hereby certify under 37 CFR §1.8(a) that this correspondence is being deposited with the United States Postal Service as first class mail with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

October 1, 2003
Date of Deposit

Teri Barnett
Signature

Teri Barnett
Typed or Printed Name of Person Signing Certificate

4) Gadi Gazit was a graduate student in my laboratory. Under my direct supervision he performed routine biochemical experiments to test the Grp78 promoter in tissue culture cells. He also followed the growth of xenografts derived from the same tissue culture cells. While he contributed to the technical aspects of the work, he did not contribute to the conception of the subject invention.

5) Gene Hung was a post-doctoral research associate in the laboratory of Dr. French Anderson. Under my direct supervision he constructed the retroviral vectors containing the Grp78 promoter driving the HSVTK suicide gene. He also performed staining of tumor tissues. Therefore, while he provided useful reagents in reducing the subject invention to practice, he did not contribute to the conception of the subject invention.

6) Xinke Chen was a post-doctoral research associate in my laboratory. Under my direct supervision he performed routine biochemical experiments to test the Grp78 promoter in tissue culture cells. While he contributed to the technical aspects of the work, he did not contribute to the conception of the subject invention.

7) W. French Anderson is a Professor in the department of Biochemistry and Molecular Biology at USC School of Medicine and was the laboratory supervisor of Dr. Hung. Dr. Anderson provided retroviral vectors which were modified under my direction. He did not contribute to the conception of the subject invention.

8) In summary, Gadi Gazit, Gene Hung, Xinke Chen and W. French Anderson, although named as co-authors on the Gazit et al. article, are not co-inventors of the subject invention but rather were persons who either worked under my direction and supervision or in a collaboration provided me with routine reagents to reduce the subject invention to practice.

I hereby declare that all statement made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of inventor: Amy S. Lee
Inventor's signature: Amy S. Lee
Date: Sept. 2, 2003

MPR/mpr

Fish & Richardson P.C.
4350 La Jolla Village Drive, Suite 500
San Diego, California 92122
Telephone: (858) 678-5070
Facsimile: (858) 678-5099

VOLUME 7 NUMBER 4 MAY 1988

ISSN 0198 0238

EXHIBIT A

DNA

A Journal of Molecular and Cellular Biology

Mary Ann Liebert, Inc. publishers

DNA
Volume 7, Number 4, 1988
Mary Ann Liebert, Inc., Publishers
Pp. 275–286

Human Gene Encoding the 78,000-Dalton Glucose-Regulated Protein and Its Pseudogene: Structure, Conservation, and Regulation

JERRY TING and AMY S. LEE

ABSTRACT

The isolation and characterization of a human functional GRP78 gene and a processed pseudogene are described. We present the complete primary structure of the human GRP78 gene, which spans over 5 kb and consists of eight exons. Sequence comparisons reveal that the GRP78 gene shares unusual homology among the human, rat, and hamster in the protein-coding and 3′ untranslated regions. In addition, short domains highly conserved with HSP70 isolated from human, *Drosophila*, *Xenopus*, yeast, and *E. coli* DNA are identified within the hydrophobic regions of GRP78. The intronless pseudogene resembles that of a processed gene. It is flanked by a short direct repeat and is embedded within an AT-rich genomic region. The highly active promoter from the functional human GRP78 gene contains a TATA box, five CCAAT sequences, and two potential binding sites for the transcriptional factor Sp1. It consists of a distal domain that enhances basal level expression and a proximal domain essential for responses to calcium ionophore and for a temperature-sensitive mutation which induce the GRP78 gene. Both domains are highly conserved between the rat and the human GRP78 promoters.

INTRODUCTION

THE GLUCOSE-REGULATED PROTEINS (GRPs) were first observed in fibroblasts following transformation with avian sarcoma viruses. Subsequently, it was found that the increased synthesis of these proteins was primarily due to the depletion of glucose and a variety of physiological stress conditions can also induce the synthesis of the GRPs. Potent inducers of the GRPs include calcium ionophores, sulfhydryl-reducing reagents, and reagents which block protein glycosylation (for review, see Lee, 1987).

The most abundant GRP is a 78,000-dalton protein localized in the endoplasmic reticulum (ER) (Zala *et al.*, 1980). It shares partial amino acid sequence homology with the 70,000-dalton heat shock protein (HSP70) and was recently identified as the immunoglobulin (Ig) heavy-chain binding protein in B lymphocytes (Munro and Pelham, 1986). Evidence has since accumulated showing that GRP78 is ubiquitous and can bind to a variety of proteins processed through the ER (Gething *et al.*, 1986; Hendershot, Bole, and Kearney, in prep.; Nakaki, Deans, and Lee, in prep.). Therefore, one plausible function of GRP78 is that under normal conditions, it may be part of a protein-processing machinery in the ER (Bole *et al.*, 1986); during stress, it may bind to aberrant or underglycosylated proteins and prevent their aggregation (Pelham, 1986).

To understand the structure and regulation of the GRP genes, the cDNA clones encoding GRP78 and GRP94 have been isolated from the hamster (Lee *et al.*, 1981). These molecular probes were used to establish that the genes encoding the GRP78 and GRP94 were regulated coordinately at the transcriptional level during glucose starvation, calcium ionophore treatment, and in a temperature-sensitive (*ts*) hamster cell line K12 which is blocked in protein glycosylation at 40°C (Lee *et al.*, 1983, 1986; Lin and Lee, 1984; Resendez *et al.*, 1985). The cloning and function of the rat GRP78 promoter have also been described (Atte-

Department of Biochemistry and the Norris Cancer Research Institute, University of Southern California School of Medicine, 1441 Eastlake Avenue, Los Angeles, CA 90033.

nello and Lee, 1984; Chang *et al.*, 1987). The rat GRP78 promoter contains a highly active enhancer and is capable of conferring glucose starvation and calcium ionophore A23187 induction upon heterologous genes (Attenello and Lee, 1984; Resendez *et al.*, 1985; Lin *et al.*, 1986). Further, transfection of these fusion genes into the K12 *ts* mutant cell line results in their induction at 40°C (Attenello and Lee, 1984; Chang *et al.*, 1987). These combined results strongly indicate the existence of *cis*-acting elements in the GRP78 promoter responsive to specific stress conditions.

A powerful approach to identify important functional domains in the GRP78 promoter and coding sequences is through direct comparisons of the GRP78 and other related genes derived from different species and searching for evolutionary conserved regions. Besides GRP78, the gene encoding hsc73, the major HSP70 cognate protein in unstressed rat cells has been isolated from the rat genome (Sorger and Pelham, 1987). The isolation of the human heat-inducible HSP70 gene has been described (Wu *et al.*, 1985). In this report, we describe the isolation and determination of the complete primary structure of the human GRP78 gene. Our analysis reveals the existence of a functional gene and a processed pseudogene of GRP78 in the human genome. The functional gene contains eight exons and shares unusual sequence homology with the rat and hamster GRP78 gene. In addition, short domains highly conserved with HSP70 isolated from human, *Drosophila*, *Xenopus*, yeast and *E. coli* DNA are identified within the hydrophobic regions of GRP78. The pseudogene resembles that of a processed gene and is embedded within an AT-rich genomic region. The promoter from the functional GRP78 gene contains a TATA box and five CCAAT sequences. It consists of a distal domain which enhances basal level expression and a proximal domain essential for responses to various stimuli. Both domains are highly conserved between the rat and the human GRP78 promoter.

MATERIALS AND METHODS

Cell lines and culture conditions

The temperature sensitive (*ts*) Chinese hamster fibroblast cell line K12 has been described (Roscoe *et al.*, 1973; Lee, 1981). It was routinely maintained in DMEM (4.5 mg/ml glucose) supplemented with 10% cadet calf serum. The HeLa D98 AH2 monolayer cells were obtained from R.E.K. Fournier (USC) and were maintained in DMEM with 10% fetal calf serum.

Treatment conditions

HeLa cells grown in 150-mm diameter dishes to 90% confluency were changed to fresh medium. The cells were either treated with 7 μM A23187 for 5 hr (Resendez *et al.*, 1985) or grown in glucose-depleted DMEM for 16 hr (Lin and Lee, 1984) prior to extraction of cytoplasmic RNA.

Isolation of human GRP78 genes

The human fetal liver genomic library (Lawn *et al.*, 1978) was screened using a full-length cDNA plasmid p3C5 which encodes the hamster GRP78 (Lee *et al.*, 1981). Filters containing about 10^6 recombinant phage were prehybridized in 50% formamide, 5× SSC (1× SSC: 0.15 M NaCl, 15 mM sodium citrate), 5× Denhardt (1× Denhardt: 0.02% each of bovine serum albumin, polyvinylpyrrolidone, and Ficoll), 0.1% $NaDodSO_4$, 50 mM sodium phosphate buffer pH 6.8, 1% glycine, and 100 μg/ml denatured salmon sperm DNA at 42°C for 1 hr. They were then hybridized in the same buffer with nick-translated p3C5 probe (sp. act. 7×10^7 cpm/μg; total cpm used 1.5×10^8) overnight at 42°C. After hybridization, the filters were washed three times at 50°C for 60 min each in 5× Denhardt, 3× SSC, 0.1% $NaDodSO_4$, and 0.1% sodium pyrophosphate; and 2 times at 50°C for 60 min each in 1× SSC, 0.1% $NaDodSO_4$, and 0.1% sodium pyrophosphate. Six recombinant phage hybridized positively with the p3C5 probe.

S_1 nuclease protection

Cytoplasmic RNA was isolated from control and induced HeLa cells as described previously (Resendez *et al.*, 1985). The DNA probe used for the S_1 protection assay was a 497-nucleotide *Sma* I fragment corresponding to nucleotides −368 to 129 (Figs. 1 and 4). The DNA was 5′-end-labeled by T4 polynucleotide kinase (BRL) to a specific activity 6×10^6 cpm/μg of DNA. About 4×10^5 cpm of the heat-denatured DNA was hybridized with 30 μg of HeLa RNA in 80% formamide, 0.4 M NaCl, 40 mM PIPES [piperazine-N,N'-bis(2-ethanesulfonic acid)] pH 6.4, and 1 mM EDTA at 58°C for 16 hr. The DNA–RNA hybrid was then digested with 20 units of S_1 nuclease (PL Biochemicals) in 200 μl of reaction mixture containing 0.28 M NaCl, 50 mM sodium acetate pH 4.6, 4.5 mM $ZnSO_4$, and 50 μg/ml of denatured salmon sperm DNA at 37°C for 30 min. Undigested DNA was precipitated with alcohol and electrophoresed on a 6% polyacrylamide sequencing gel.

Primer extension

The primer used was a synthetic oligonucleotide (a 17-mer, with sequence 5′-GGCCGCGACGCTTACCT-3′) corresponding to anti-sense nucleotides 24–40 of the rat GRP78 5′ UTR (Chang *et al.*, 1987). The primer was labeled by T4 polynucleotide kinase to a specific activity of 1.5×10^8 cpm/μg DNA. About 10^6 cpm of the primer was hybridized with 50 μg of HeLa RNA. The hybridization and extension reactions were performed as described (Lin *et al.*, 1986), with the exception that hybridization was performed at 25°C for 5 hr in the same hybridization buffer as in S_1 nuclease protection experiment.

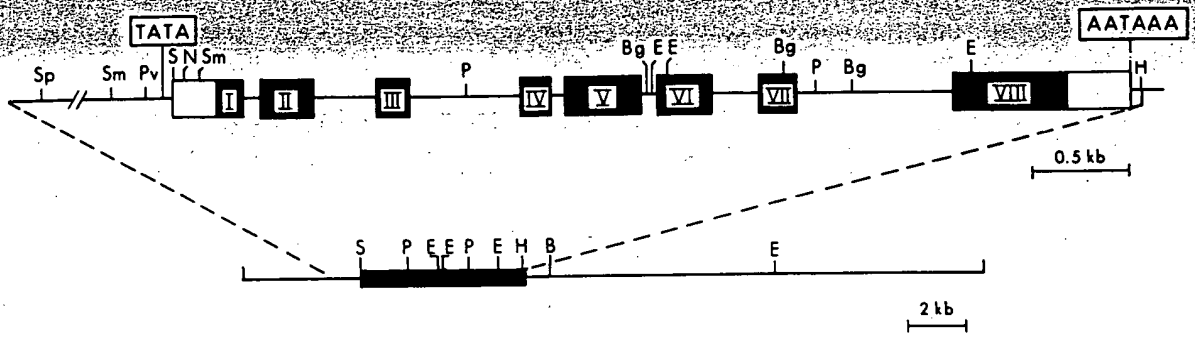


FIG. 1. Restriction map and exon structure of the GRP78 gene from clone hu28-1. The lower line shows the restriction map of the 24-kb human genomic DNA within the phage. The upper line shows the expanded exon structure of the GRP78 gene, with the 5′ and 3′ untranslated regions as open boxes, and the coding region in black, numbered boxes. The TATA box and the poly(A) addition sequence AATAAA are indicated. Restriction enzyme sites marked are: *Bam* HI(B), *Bgl* II (Bg), *Eco* RI (E), *Hind* III (H), *Nru* I (N), *Pst* I (P), *Pvu* II (Pv), *Sal* I(S), *Sma* I (Sm), and *Sph* I (Sp).

Construction of hybrid genes

To construct plasmids phu78CAT (−1650), phu78CAT (−368), and phu78CAT (−170), DNA fragments *Sph* I-*Nru* I (−1650 to 53), *Sma* I-*Sma* I (−368 to 129), or *Pvu* II-*Nru* I (−170 to 53) containing the human GRP78 promoter and 5′ UTR sequences (Fig. 1) were blunt-ended with T4 DNA polymerase and ligated into the unique *Hind* III site (blunt-ended with T4 DNA polymerase) of the pSVOCAT vector (Resendez *et al.*, 1985). Plasmids containing the GRP78 promoter and CAT gene fused in the same transcriptional orientation were identified by restriction mapping.

Transient transfection conditions and CAT assay

Transfection of DNA into K12 cells and the assays for the chloramphenicol acetyltransferase (CAT) activity have been described (Resendez *et al.*, 1985).

DNA sequence analysis

The DNA contained within the phage insert was restriction-mapped. Purified DNA fragments were subcloned into M13 vector and sequenced by the dideoxy chain-termination method (Sanger *et al.*, 1977). Most of the DNA fragments were sequenced on both strands.

Computer programs

Analysis of the DNA sequences was aided by the Intelligenetics Bionet program. The protein hydropathicity was analyzed by using the University of Wisconsin Genetics Computer Group (UWGGG) program.

RESULTS

Isolation of two types of human GRP78 genes

A human genomic library was screened with a full-length cDNA probe encoding the hamster GRP78 transcriptional unit. Six positive plaques were isolated; they represented five different λ clones based on limited restriction mapping. The clones can be further grouped into two restriction patterns, indicating that two different GRP78 genes derived from different regions of the human genome have been isolated. The recombinant phage, hu28-1 and hu3-2B, representing the two types, were selected for detailed restriction mapping and probes derived from the hamster cDNA were used to identify the limits of the coding sequences in the λ insert (Figs. 1 and 2). The hu28-1 clone contains a GRP78 gene spanning 5 kb interrupted by intervening sequences, whereas hu3-2B contains a contiguous coding region within 2 kb.

The structure of the functional GRP78 gene

To determine if the GRP78 gene contained within hu28-1 is functional, we searched for canonical promoter and transcriptional initiation and termination elements at the 5′ and 3′ borders of this gene. The transcriptional start site was mapped by both S_1 nuclease protection assay and primer extension (Fig. 3A). For this purpose, total cytoplasmic RNA was isolated from HeLa cells induced by calcium ionophore A23187 or glucose starvation. As a control, RNA was also extracted from noninduced cells. In the S_1 assay, RNA samples were hybridized with a 497-nucleotide *Sma* I fragment containing the 5′ border of the GRP78 gene. As shown in Fig. 3B, the major S_1-resistant band has a size of 129 nucleotides. Using the primer extension method, a discrete band of 80 nucleotides corresponding to the site mapped by S_1 protection was detected in the induced RNA sample (Fig. 3C). The smaller inducible prod-

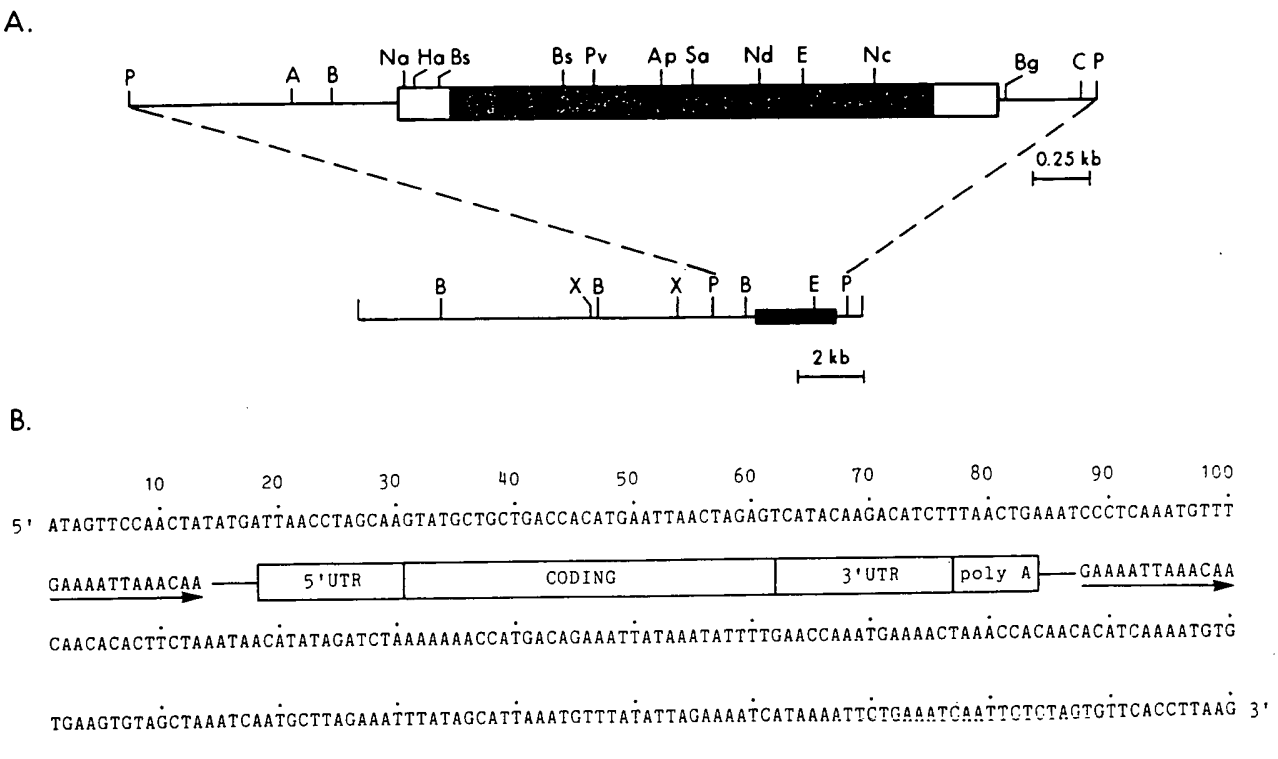


FIG. 2. The processed pseudogene from clone hu 3-2B and the sequence surrounding the insertion site. A. The lower line shows the restriction map of the 15-kb human genomic DNA insert within the phage. The upper line shows the expanded region spanning the 4.4-kb *Pst* I fragment containing GRP78 coding sequences (black box) and untranslated regions (open boxes). The restriction sites marked are *Ava* I (A), *Apa* I (Ap), *Bam* HI (B), *Bgl* II (Bg), *Bst* EII (Bs), *Cla* I (C), *Eco* RI (E), *Hae* II (Ha), *Nae* I (Na), *Nco* I (Nc), *Nde* I (Nd), *Pvu* II (Pv), *Pst* I (P), *Sac* I (Sa), *Xba* I (X). B. The nucleotide sequences immediately surrounding the pseudogene. The 13-nucleotide direct repeat sequence flanking the inserted gene is marked by arrows.

uct (63 nucleotides) is probably due to premature stop of the reverse transcriptase. These combined results identified the transcriptional initiation site as 129 nucleotides upstream from the 3 *Sma* I site. Sequencing the DNA 5′ of the cap site revealed a TATA sequence 25 nucleotides upstream. In addition, there are five CCAAT sequences located within 250 nucleotides upstream from the RNA initiation site. Two potential Sp1 binding sites were localized at −228 and −164. At the 3′ end of the gene, a polyadenylation sequence AATAAA was also found. All of these features resemble that of a functional gene.

To understand in detail the organization of the GRP78 gene, the complete primary structure of the human GRP78 gene contained in the hu28-1 clone was determined by DNA sequencing. This sequence data, combined with the previously obtained full-length cDNA sequence of the rat and hamster GRP78 (Munro and Pelham, 1986; Ting *et al.*, 1987), enabled us to delineate the locations of exon-intron junctions. As shown in Figs. 1 and 4, clone hu28-1 contains a gene comprising the entire GRP78 coding sequence interrupted by seven intervening sequences. All the exon-intron junctions followed the GT-AG splicing rule (Breathnach *et al.*, 1978). Comparing this gene with the rat GRP78 gene (S.K. Wooden, R.P. Kapur, and A.S. Lee, in prep.), the number and location of the exons are identical. However, the sizes of the intervening sequences are similar but not identical between the rat and human.

GRP78 protein contains highly conserved domains

The complete amino acid sequence was deduced from the determined nucleotide sequence and was subjected to the hydropathicity analysis (Kyte and Doolittle, 1982). As shown in Fig. 5, this protein is very hydrophilic, especially toward the carboxyl half of the protein. Only a few small regions were identified to be hydrophobic. Upon comparison of the entire amino acid sequences derived from the human GRP78, rat GRP78, and HSP70 from human, *Drosophila*, *Xenopus*, yeast and *E. coli dnaK* gene, a few regions (A–H) of high sequence homology (over 80%) can be detected (Fig. 5). Similar conservations were observed with the rat hsc73. Strikingly, these domains cover most of the hydrophobic regions of GRP78, including the previously reported 11-amino-acid domain (A) near the amino terminus (Chappell *et al.*, 1986; Chang *et al.*, 1987). The

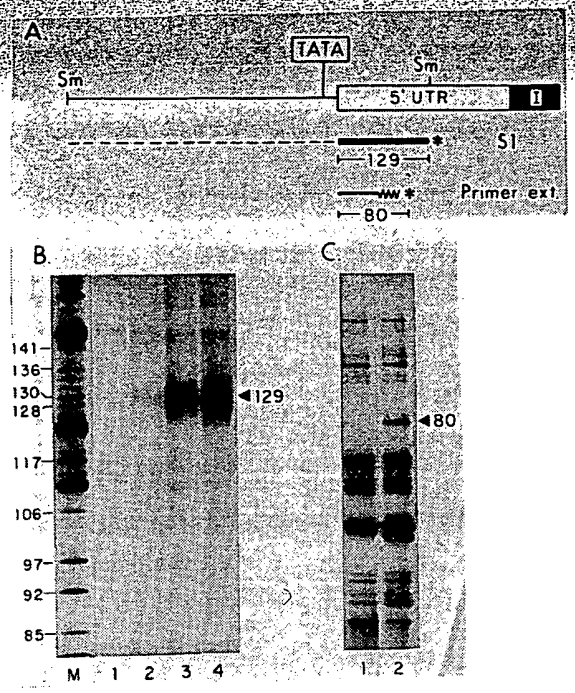


FIG. 3. Determination of the transcriptional start site. A. Schematic representation of the probes used for S_1 mapping and primer extension. The asterisks indicate the kinase-labeled termini. B. S_1 mapping of GRP78 transcripts in HeLa cells. RNA samples were from yeast tRNA (lane 1), cells grown continuously in DMEM (lane 2), cells grown in glucose-depleted DMEM for 16 hr (lane 3), and cells grown in DMEM with 7 μM A23187 for 5 hr (lane 4). The lane marked M refers to a size marker from a dideoxy sequencing reaction run in parallel on the same gel. C. Primer extension using a synthetic primer. The RNA samples isolated from HeLa cells grown in the absence (lane 1) or presence of A23187 (lane 2) were used. The resultant DNA fragments were electrophoresed on a 6% polyacrylamide sequencing gel with DNA fragment size markers. The autoradiograms are shown. The sizes (nucleotides) of the major S_1 protected band and the primer extension product are indicated.

other six domains fall into two symmetrical groups of three (B,C,D and E,F,G) around the center of the protein. The most divergent sequences were found at the highly hydrophilic region near the carboxyl third of the protein. Interestingly, the last four amino acids (Lys-Asp-Glu-Leu) which are important for retaining GRP78 in the endoplasmic reticulum (Munro and Pelham, 1987) are conserved among the human, rat, and hamster GrP78 and are lost from the cytoplasmic HSP70.

Structure of the processed human GRP78 gene

In contrast to hu28-1, recombinant phage hu3-2B does not contain any introns. The coding region was mapped by hybridization to the hamster DNA and by S_1 nuclease digestion, using RNA extracted from HeLa cells (data not shown). These analyses, coupled with sequencing of the DNA flanking the coding region, revealed that this recombinant contains features characteristic of a processed gene (Fig. 2B). First, the coding region is contiguous throughout the gene (Ting, 1988). When compared to the functional gene, we detected a 12-nucleotide deletion close to the amino terminus of the mature protein and a 4-nucleotide insertion within exon VI which results in a shift of the reading frame. In addition, point mutations were found throughout the coding region. At the 3′ end of the gene, the consensus sequence for the poly(A) addition site AATAAA was present; shortly downstream from that sequence, there was a stretch of adenosine residues. These results clearly indicate that this gene was derived from a RNA intermediate of the GRP78 gene.

As in the case of other processed genes (Hollis *et al.*, 1982; Wilde *et al.*, 1982; Karin and Richards, 1982), a short direct repeat sequence, which probably represents the site of insertion of the processed gene, was located flanking the inserted sequence. The repeat is an AT-rich 13-mer with the sequence GAAAATTAAACAA. The sequence immediately surrounding the processed gene is strikingly AT rich. The 100-nucleotides upstream from the inserted gene are 66% AT and the 200 nucleotides downstream from the inserted gene are 75% AT. When a 0.9-kb fragment containing the 5′-flanking region of this processed gene was fused to bacterial CAT gene, we could not detect any CAT activity after transient transfection into K12 cells (Ting, 1988).

Expression and regulation of the human GRP78 promoter

In hamster cells, the GRP78 gene is transcriptionally activated by the calcium ionophore A23187 (Resendez *et al.*, 1985) and the K12 *ts* mutation (Lee *et al.*, 1983). By deletion and *in vivo* competition, the important region for these responses in the rat GRP78 promoter is mapped within 430 nucleotides upstream from TATA box (Lin *et al.*, 1986; Chang *et al.*, 1987). To test if the human GRP78 gene isolated from the hu28-1 clone contains a functional promoter inducible by A23187 and the K12 *ts* mutation, human GRP78 promoter/CAT fusion genes were constructed and transfected into K12 cells. As a comparison, the CAT fusion gene pI10 containing the rat GRP78 promoter (Resendez *et al.*, 1985) was also tested. After 30 hr, the cells were treated with 7 μM A23187 or shifted to the nonpermissive temperature 40°C. Cell extracts were assayed for the CAT activity (Fig. 6A) and the promoter activities of the CRP78/CAT fusion genes under induced or non-induced conditions are summarized (Fig. 6B).

Our results show that the human GRP78 promoter is capable of conferring A23187 and K12 *ts* mutation induction to the CAT gene. Further, no significant changes were observed at basal or induced activity when the 5′-flanking sequence of the human GRP78 gene was reduced from

```
CCCGGGGTCACTCCTGCTGGACCTACTCCGACCCCCTAGGCCGGGAGTGAAGGCGGGACTTGTGCGGTTACCAGCGGAAATGCCTCGGGGTCAGAAGTCGCAGGAGAGATAGACAGCTGC  120

TGAACCAATGGGACCAGCGGATGGGGCGGATGTTATCTACCATTGGTGAACGTTAGAAAC GAATAGCAGCCAATGAATCAGCTGGGGGGGCGGAGCAGTGACGTTTATTGCGGAGGGGGC  240

CGCTTCGAATCGGCGGCGGCCAGCTTGGTGGCCTGGGCCAATGAACGGCCTCCAACGAGC AGGGCCTTCACCAATCGGCGGCCTCCACGACGGGGCTGGGGGAGGGTATATAAGCCGAGT  360

AGGCGACGGTGAGGTCGACGCCGGCCAAGACAGCACAGACAGATTGACCTATTGGGGTGTTTCGCGAGTGTGAGAGGGAAGCGCCGCGGCCTGTATTTCTAGACCTGCCCTTCGCCTGGT  480

                                                                                                                    M  K  L
TCGTGGCGCCTTGTGACCCCGGGCCCCTGCCGCCTGCAAGTCGAAATTGCGCTGTGCTCC TGTGCTACGGCCTGTGGCTGGACTGCCTGCTGCTGCCCAACTGGCTGGCAAGATGAAGCT  600

  S  L  V  A  A  M  L  L  L  L  S  A  A  R  A  E  E  E  D  K  K  E  D  V  G  T  V  V  G  I  D  L  G  T  T  Y  S  C<---IN-
CTCCCTGGTGGCCGCGATGCTGCTGCTGCTCAGCGCGGCGCGGGCCGAGGAGGAGGACAA GAAGGAGGACGTGGGCACGGTGGTCGGCATCGACTTGGGGACCACCTACTCCTGGTAAGT  720

TRON 1                                                                                                  INTRON 1---> V  G  V  F  K  N  G  R  V  E  I  I
GGGGTTGCGGATGAGGGGGACGGGGCGTGGCGCTGGCTGGCGTGAGAAGTGCGGTGCTGA TGTCCCTCTGTCGGGTTTTTGCAGCGTCGGCGTGTTCAAGAACGGCCGCGTGGAGATCAT  840

  A  N  D  Q  G  N  R  I  T  P  S  Y  V  A  F  T  P  E  G  E  R  L  I  G  D  A  A  K  N  Q  L  T  S  N  P  E  N  T  V  F
CGCCAACGATCAGGGCAACCGCATCACGCCGTCCTATGTCGCCTTCACTCCTGAAGGGGA ACGTCTGATTGGCGATGCCGCCAAGAACCAGCTCACCTCCAACCCCGAGAACACGGTCTT  960

  D  A  K  R  L  I  G  R  T  W  N  D  P  S  V  Q  Q  D  I  K  F  L  P  F  K <----INTRON 2
TGACGCCAAGCGGCTCATCGGCCGCACGTGGAATGACCCGTCTGTGCAGCAGGACATCAA GTTCTTGCCGTTCAAGGTTCGACCGGTTTTCCTCATCCAGTTAGAGAACGGGTGGGTGGT 1080

GGGAGTATTTAGAGTTATAAGTCTCTGGAAAAGTGTTGAGACAACAGTTGAAGGTTATAG ACATGATGTATGTAATAACTTTAATACTATTAGTATGTTACAAAACTTAAGACAGTTGCT 1200

GTCGTACTGTCTACGATAGTTTAGGAATAAAAGACCGATTAAAACTGAACTTTGTAAGACACCTATACTCCCTGAAGTATTTCTAGTCAATTTGCAGCCCCAAGGGACCAAAATAAACCA 1320

                                                                                                  INTRON 2---> V  V  E  K  K  T  K  P  Y  I  Q  V  D
AATTGTGGGGATGGTAGTGGGTCTTTTAAACTTTGAGATGTCATTGTATCTGTGTCTGAA AACAATAATTCTTTAAAATAGGTGGTTGAAAAGAAAACTAAACCATACATTCAAGTTGAT 1440

  I  G  G  G  Q  T  K  T  F  A  P  E  E  I  S  A  M  V  L  T  K  M  K  E  T  A  E  A  Y  L  G  K  K <----INTRON 3
ATTGGAGGTGGGCAAACAAAGACATTTGCTCCTGAAGAAATTTCTGCCATGGTTCTCACT AAAATGAAAGAAACCGCTGAGGCTTATTTGGGAAAGAAGGTAAATATTTCTAGAACAATG 1560

TTAAGTATTTTTTGATCATTAGTATTCTCGGTTGGCTGTTATGTATAGAAGCCTTCGTGA AGGGTTTCAAAAATTTTAATCAGAATGGTATTCATGCTTGTCACGGTTTAATTATTGAGT 1680

CCCTTTACTATAAGCCAAACAAAAATAGACTTTTCATGTATTATTTAATGCTTACAATTCCAGGAACAATAAAATTTTATATGTTGTATTCATCAATAATTGGCTTAAAAACTAAAGTGA 1800

TGGTTTGACTGTAATTTTTTTTTTTTTGAGATGGAGTCTTGCTCTGTTGCCCAGGCTGGACTGCAGTGGCACGATCTCAGCTCACTGCAACCTCTGCCTCCCGGGTTAAGCAGCTCTCCTG 1920

CCTCAGCCTCCAAGTAATGGAACGACAGGCACACCACCACAGCTGGCTAATTTTTTTTTTTTTTTTTAATTTTCAGTAGAGACAGGGTTTCTCCACATTGCCAGGCTGGTCTTGAAATCC 2040

TGCCCTCAGGTTGATCCTCCTGCCTAGCCTCCCAAAGTGCTGGATTATAGGCAGAAGCCACCGCCTGGCCAGACTGTAATTTAAATAAGGGTTAAACTATGTGACAATACACTTAATTAT 2160

  INTRON 3---> V  T  H  A  V  V  T  V  P  A  Y  F  N  D  A  Q  R  Q  A  T  K  D  A  G  T  I  A  G  L  N  V  M  R  I  I
CTTTATCCTTTTAGGTTACCCATGCAGTTGTTACTGTACCAGCCTATTTTAATGATGCCCAACGCCAAGCAACCAAAGACGCTGGAACTATTGCTGGCCTAAATGTTATGAGGATCATCA 2280

                                                                                                  INTRON 4---> T  A  A  A  I  A  Y  G
N  E  P<---INTRON 4
ACGAGCCGTAAGTATGAAATTCAGGGATACGGCATATTTGCCAAATAGTGGAAATGTGAAGTACTGACAAAACTTTTCCCTTTTTCAATCTAATAGTACGGCAGCTGCTATTGCTTATGG 2400

  L  D  K  R  E  G  E  K  N  I  L  V  F  D  L  G  G  G  T  F  D  V  S  L  L  T  I  D  N  G  V  F  E  V  V  A  T  N  G  D
CCTGGATAAGAGGGAGGGGGAGAAGAACATCCTGGTGTTTGACCTGGGTGGCGGAACCTTCGATGTGTCTCTTCTCACCATTGACAATGGTGTCTTCGAAGTTGTGGCCACTAATGGAGA 2520

  T  H  L  G  G  E  D  F  D  Q  R  V  M  E  H  F  I  K  L  Y  K  K  K  T  G  K  D  V  R  K  D  N  R  A  V  Q  K  L  R  R
TACTCATCTGGGTGGAGAAGACTTTGACCAGCGTGTCATGGAACACTTCATCAAACTGTACAAAAAGAAGACGGGCAAAGATGTCAGGAAGGACAATAGAGCTGTGCAGAAACTCCGGCG 2640

  E  V  E  K  A  K  A  L  S  S  Q  H  Q  A  R  I  E  I  E  S  F  Y  E  G  E  D  F  S  E  T  L  T  R  A  K  F  E  E  L  N
CGAGGTAGAAAAGGCCAAGGCCCTGTCTTCTCAGCATCAAGCAAGAATTGAAATTGAGTCCTTCTATGAAGGAGAAGACTTTTCTGAGACCCTGACTCGGGCCAAATTTGAAGAGCTCAA 2760

                                                                                                                    INTRON 5---> D
M <---INTRON 5
CATGGTATGTTCCTTGTTTTCTGCTTTGCTAATGAGATCTCCTTAGACTCTGAATTCAGGACATTGCATCTAGATACTTAGATAACAGACATCACAGTAACCATGTCTTTTTTCTAGGAT 2880

  L  F  R  S  T  M  K  P  V  Q  K  V  L  E  D  S  D  L  K  K  S  D  I  D  E  I  V  L  V  G  G  S  T  R  I  P  K  I  Q  Q
CTGTTCCGGTCTACTATGAAGCCCGTCCAGAAAGTGTTGGAAGATTCTGATTTGAAGAAGTCTGATATTGATGAAATTGTTCTTGTTGGTGGCTCGACTCGAATTCCAAAGATTCAGCAA 3000

  L  V  K  E  F  F  N  G  K  E  P  S  R  G  I  N  P  D  E  A  V  A  Y  G  A  A  V  Q  A  G  V  L  S  G  D  Q  D  T <---IN-
CTGGTTAAAGAGTTCTTCAATGGCAAGGAACCATCCCGTGGCATAAACCCAGATGAAGCTGTAGCGTATGGTGCTGCTGTCCAGGCTGGTGTGCTCTCTGGTGATCAAGATACAGGTAGG 3120

 TRON 6
TCATCATCGCAGCATCTTTCTTAGTGATTCAGTAGCTTGATGGAAGAGCTCGGTACCCCTATTGCTTTAGAAAATACCAGAATATGAGCAACAAGGTCACACAGCTAGTAAAGGGTATAA 3240

GTGAAGACAAGACTGGGGTAGTCTCCAAGATCATTAGCAACTGTTTAATTCACTGCCTTTAAAATGTGTGTGTTAGAACCTAACCAAATGTTAGAGAGATAAACTTTACATAGCTCATAG 3360
```

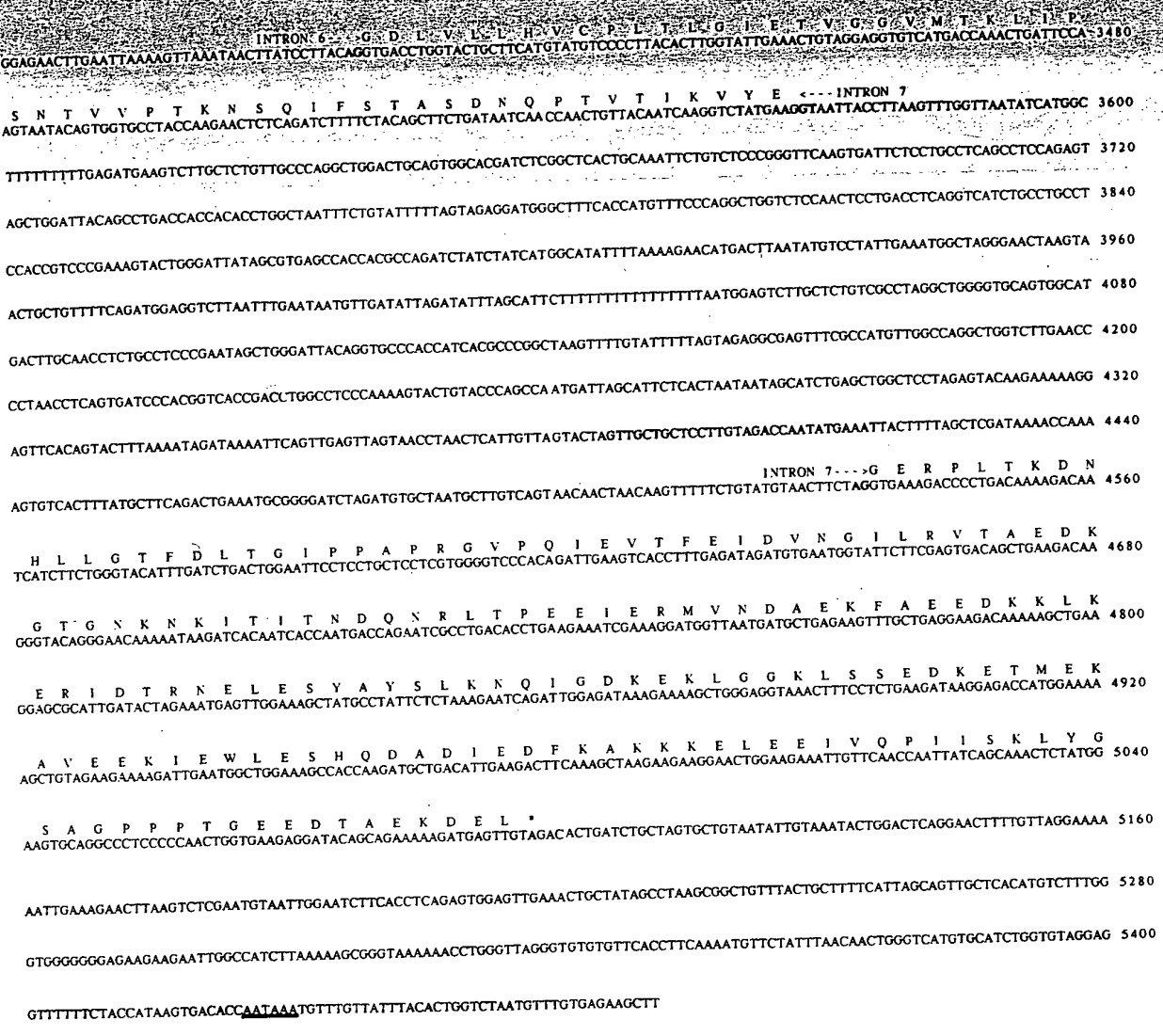
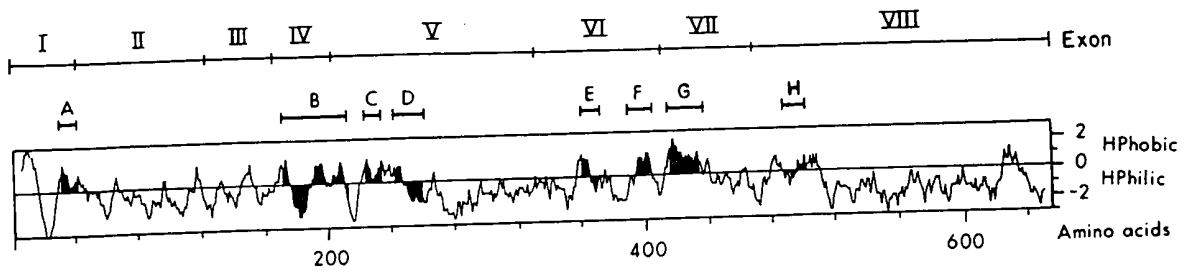


FIG. 4. Primary structure of the human GRP78 gene. The DNA sequence is determined from λ clone hu28-1. The intron–exon boundaries were mapped by comparison with the rat and hamster cDNA sequences (Ting *et al.*, 1987). The GT-AG exon–intron boundaries are indicated by boldface. The TATA and poly(A) addition sequences are underlined. The bent arrow indicates the major transcriptional start site. The termination codon TAG is indicated by a star. The amino acid sequence is shown in the one-letter code. Black triangle indicates the proteolytic cleavage site that excises the 18-amino-acid leader sequence and produces the amino terminus of the mature protein.

1650 nucleotides to 368 nucleotides (our unpublished results and Fig. 6B). When only 170 nucleotides were included, there was a twofold decline in both basal and induced activities. However, all three constructs showed about the same fold of induction at 40°C and by A23187. On the basis of these results, we conclude that the human GRP78 promoter contains a stress-inducible region that responds to A23187 or K12 *ts* mutation contained within 170 nucleotides upstream from transcriptional initiation site and a distal region at −368 to −170 which enhances the basal level of expression about twofold.

Promoter sequence comparison between human, rat, and chick GRP genes

Comparison of the rat and human GRP78 promoter sequences revealed a strong homology (78%) within 340 nucleotides upstream from the RNA initiation site of the human GRP78 gene (Fig. 7A). Several interesting sequences found in the rat GRP78 promoter (Lin *et al.*, 1986) including four CCAAT boxes, and a putative Sp1 binding site are conserved in the human GRP78 promoter. The proximal 170 nucleotides found to be important for the induc-



DOMAIN (A) GIDLGTTYSCV
DOMAIN (B) TVPAYFNDAQRQATKDAGTIAGLNVMRIINEPTAAAIAYGLD
DOMAIN (C) DLGGGTFDVS
DOMAIN (D) FEVVATNGDTHLGGEDFD
DOMAIN (E) VLVGGSTRIPKIQ
DOMAIN (F) NPDEAVAYGAAVQ
DOMAIN (G) LLHVCPLTLGIETVGGVMT
DOMAIN (H) GVPQIEVTF

FIG. 5. Conservation of amino acid domains in the human GRP78. The hydropathicity profile of the human GRP78 is aligned with the exon boundaries and the amino acid sequence. The shaded regions correspond to domains (A–H) within the human GRP78 amino acid sequence which share over 80% homology with the amino acid sequence reported for the human, *Xenopus*, *Drosophila*, yeast, and *E. coli* heat-shock inducible HSP70 genes (Lindquist, 1986). The amino acid sequences of the shared domains are shown in the one-letter code.

tive responses are highly conserved (86%). The distal domain important for high basal level shares about 70% sequence homology. Notably, the rat and human GRP78 promoter sequences start to diverge outside these critical regions. Beyond −340, the sequence homology is less than 20%.

Since the GRP78 and GRP94 genes are activated simultaneously by many stress conditions (Lee, 1987), we searched for sequence elements in the GRP78 and GRP94 promoters for consensus sequences which may account for their coordinate regulation. For this purpose, the promoter sequence reported for the chick GRP94 protein (Kleinsek *et al.*, 1986; Munro and Pelham, 1987) is compared with that of the rat and human GRP78 genes. The chick GRP94 promoter sequence is divergent from the rat and human GRP78 promoters except for a region of 22 nucleotides upstream from the TATA sequence (Fig. 7B). This highly conserved region within the three promoters examined lies within the critical region required for A23187 and K12 *ts* mutation and is devoid of CCAAT and Sp1 binding sites.

DISCUSSION

We have described the isolation of two genomic clones encoding a functional and a processed human GRP78 gene. The functional gene is separated by intervening sequences into 8 exons while the processed gene is contiguous. Interestingly, among the six positive recombinant phage identified, only two different restriction map patterns emerged. These results, together with the genomic blot analysis which showed only a few bands hybridizing with the cDNA probe at stringent criterion (data not shown), suggest that there may only be one functional and one full-length processed GRP78 gene in the human genome.

The functional GRP78 gene, in contrast to intronless heat-inducible HSP70 genes isolated so far, contains a leader sequence and seven introns. Structurally, this gene resembles more the intron-containing, heat shock noninducible hsc73 gene reported in the rat (Sorger and Pelham, 1987). However, GRP78 contains a signal sequence which targets the protein into the endoplasmic reticulum and extra amino and fewer carboxyl terminal sequences as compared with hsc73. Interestingly, sequences homologous to part of the 5′ untranslated region (UTR) and the extra amino-terminal sequences of the rat GPR78 gene are found within the first additional intron of the rat hsc73 located in its 5′ UTR (S.K. Wooden, R.P. Kapur, and A.S. Lee, in prep.).

From tryptic peptide maps of the hamster and chicken GRP78, we predicted previously that it must be highly conserved through evolution (Lee *et al.*, 1981). More direct evidence came from the comparison of nucleotide sequence between hamster and rat which revealed a 94% homology in the coding region (Ting *et al.*, 1987). The magnitude of conservation between the primate gene reported here and the rodent gene reported previously is also surprisingly high. Although they diverged from each other about 75 million years (my) ago, the percent of homology in the coding DNA region is as high as 91% between these two species. Based on this conservation, we calculated that the rate of silent site mutation within the coding region is only 0.1% per my as compared to the 0.5–1% average mutation rate generally observed (Hayashida and Miyata, 1983; Yaffe *et al.*, 1985). The higher rate of generating substitution in the introns (0.4–0.5% per my) as compared

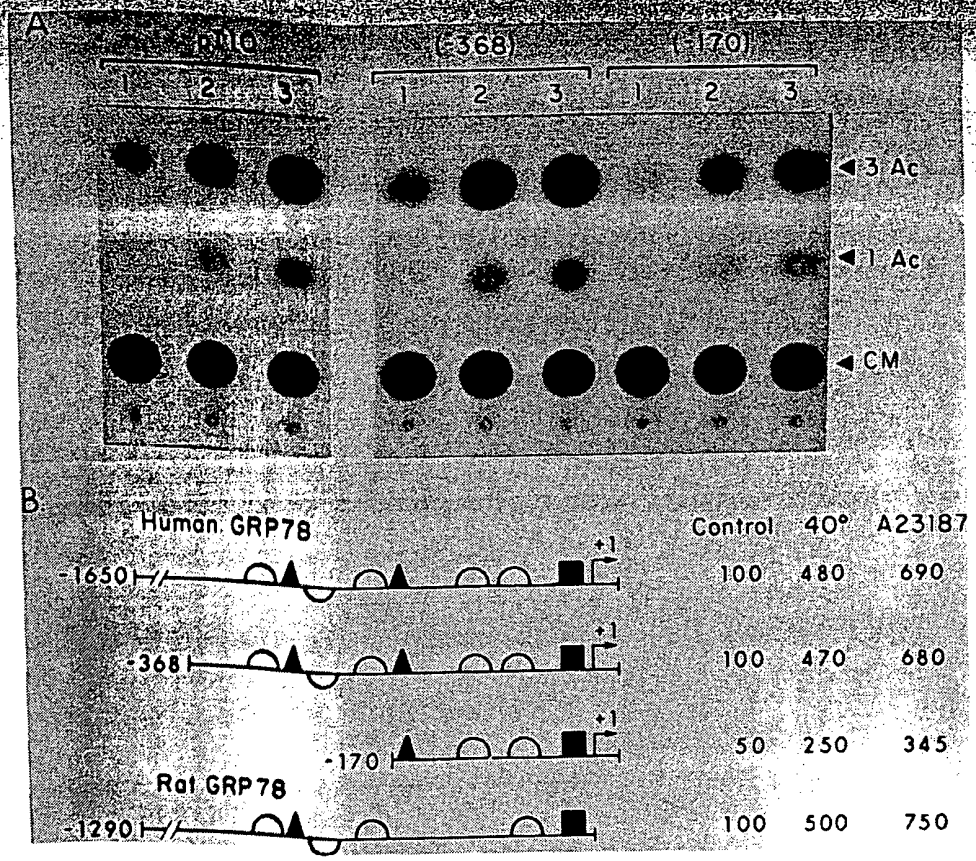
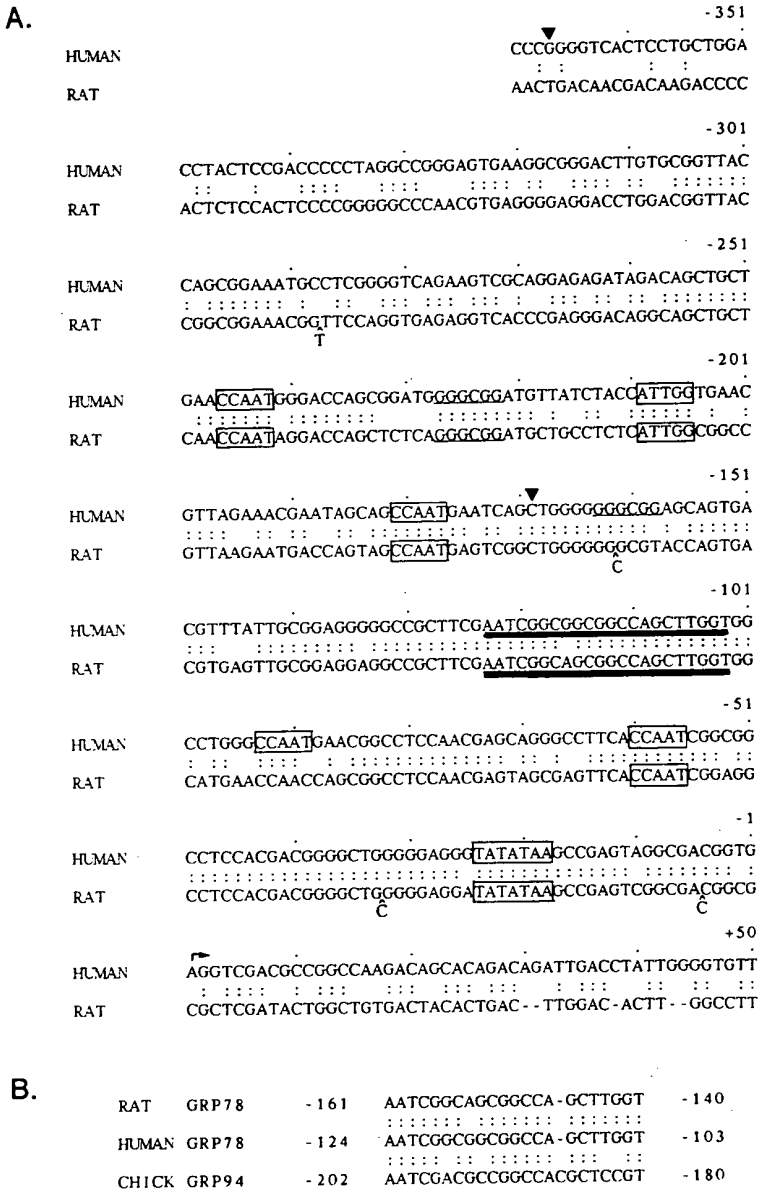


FIG. 6. Promoter activities of human GRP78-CAT fusion gene constructs. A. K12 cells were transfected with 3 μg of CAT fusion genes containing various lengths (368 and 170 nucleotides upstream from the CAP site) of the human GRP78 promoter. As a comparison, a CAT construct (pI10) which contains 1290 nucleotides of the rat GRP78 promoter (Resendez *et al.*, 1985) is also included. After 30 hr, the cells were either maintained at 35°C (lane 1), shifted to 40°C (lane 2), or treated with 7 μM A23187 at 35°C for 16 hr (lane 3). Protein extracts from the transfectants were assayed for CAT activities. The positions of chloramphenicol (CM) and its acetylated forms (3-Ac and 1-Ac) are indicated. B. Summary of the promoter activities. The promoter features of the GRP78 promoter contained within the CAT fusion gene constructs are schematically presented in the line diagrams. The symbols marked are (■) TATA, (⌓) CCAAT, (⌔) CCAAT inverted, (▲) putative Sp1 binding site, and (⌐→) the major RNA initiation site from the GRP78 gene. The results represent the average of three separate experiments and are expressed as a percentage of the activity produced by the −1650 deletion of the human GRP78 promoter in the absence of temperature shift or A23187 treatment.

with the coding region indicates that there is selective pressure to conserve the GRP78 coding region. The 3′ UTR shares 88% homology between human and rat. The conserved region covers the entire 3′ UTR even beyond the polyadenylation sequence, and the degree of divergence within the proximal half of the 3′ UTR is much lower than in the distal half. This observation is an exception to the general rule that the rate of substitution in the distal portion is about half the rate in the proximal portion of 3′ UTR (Miyata *et al.*, 1981). Several other examples also reported strong conservation in the 3′ UTR region. Genes from various mammalian species coding for the β-actin, cardiac muscle actin, the α-subunit of glycoprotein hormone, and the c-*fos* contain unusual high sequence homology at the 3′ UTR (Yaffe *et al.*, 1983; Van Beveren *et al.*, 1983; Van Straaten *et al.*, 1983). While the biological significance of conservation in the 3′ UTR is not clear, it might regulate gene expression at the translational level (Liebhaber and Kan, 1982).

With the identity of GRP78 as the immunoglobulin heavy-chain binding protein (Munro and Pelham, 1986), GRP78 represents a stress protein with a demonstrable binding function to other proteins. It may do so by binding to the hydrophobic regions of proteins improperly folded in the ER and prevent their aggregation. Since the expression of GRP78 is correlated with the inhibition of protein glycosylation (Chang *et al.*, 1987), it is possible that under stress conditions, the glycosylation process is disrupted resulting in the accumulation of underglycosylated proteins in the endoplasmic reticulum. Without the sugar moiety some proteins may not fold properly and some hydrophobic regions normally embedded in the



B.

RAT	GRP78	-161	AATCGGCAGCGGCCA-GCTTGGT	-140
HUMAN	GRP78	-124	AATCGGCGGCGGCCA-GCTTGGT	-103
CHICK	GRP94	-202	AATCGACGCCGGCCACGCTCCGT	-180

FIG. 7. Sequence comparison between the human and the rat GRP78 promoter. A. The promoter sequence of rat (Chang *et al.*, 1987) and human GRP78 gene reported here are aligned to obtain maximal homology. The sequence is numbered starting from the major RNA cap site (┌►) of the human GRP78 gene as +1. Identical nucleotides are indicated by vertical dots. The TATA and CCAAT sequences are boxed. The putative Sp1 binding sites are underlined. The deletion end points of the two GRP78-CAT fusion genes tested (−368 and −170) are indicated (▼). B. The consensus sequence found in the promoter regions of the human, rat GRP78 gene and the chicken GRP94 gene (Kleinsek *et al.*, 1986). The sequences are numbered starting from their RNA initiation site. The location of this sequence in the human and rat GRP78 promoter is marked by a heavy line in A.

hydrophilic environment or shielded by the sugar residues may now be exposed. These hydrophobic domains may be the binding sites for GRP78. In support of this hypothesis, the Ig heavy-chain constant domain C_H1, which binds to GRP78, is hydrophobic and normally covered by the light chain (Hendershot *et al.*, 1987). It is postulated that binding of GRP78 to the heavy chain prevents the self-aggregation and the premature secretion of these proteins (Bole *et al.*, 1986). To carry out this proposed function, it is desirable that GRP78 be a soluble protein. It also requires that GRP78 contains some hydrophobic domains to interact with the exposed hydrophobic regions of the improperly folded proteins. From the hydropathicity profile of GRP78, it is a highly hydrophilic protein and therefore

should be soluble. It also contains several small hydrophobic regions. Compared to the hydrophilic regions, these domains are more conserved among proteins of the HSP70 family. Since HSP70 is also postulated to have protein binding ability (Pelham, 1986; Pinhasi-Kimhi *et al.*, 1986), it is possible that these conserved regions are the critical protein binding domains within the stress gene protein family. *In vitro* mutation of the conserved domains may yield important information about their functional significance.

From the analysis of the rat GRP78 gene promoter, the regulatory elements for K12 *ts* mutation and calcium ionophore A23187 induction are localized within 430 nucleotides upstream of the TATA sequence (Chang *et al.*, 1987). A 291-nucleotide subfragment from the rat GRP78 promoter fused to heterologous promoters is capable of responding to various stimuli, as well as increasing substantially the basal level activity of the promoters (Lin *et al.*, 1986; Y.K. Kim and A.S. Lee, in prep.). This suggests that there are at least two distinct regulatory elements in the GRP78 promoter. Our deletion studies of the human GRP78 promoter further revealed that a 170-nucleotide fragment immediately upstream from the RNA initiation site contains the DNA sequence requirement for the K12 *ts* mutation and calcium ionophore induction. Further upstream sequences enhance the basal level expression. The discovery of the consensus sequence among the otherwise divergent GRP78 and GRP94 promoters would allow us to test whether this short conserved sequence is functionally significant in the *in vivo* regulation of the GRP genes by stress. Detailed analysis of the *cis*-acting DNA regulatory elements in the GRP78 promoter, coupled with the isolation of the transcriptional regulatory factors, will provide the critical information on its regulatory mechanism.

ACKNOWLEDGMENTS

We thank M.F. Chang for his assistance in the analysis of the hydropathicity profile and S.K. Wooden for stimulating discussion. J. Ting was supported in part by the California Foundation for Biochemical Research. This research was supported by Public Health Services grant CA27607 from the National Institutes of Health. A.S. Lee is a recipient of the American Cancer Society Faculty Research Award. Computer resources used to carry out our sequence analysis were provided by the Bionet National Computer Resource funded by NIH grant (1U41 RR-01695-01).

REFERENCES

ATTENELLO, J.W., and LEE, A.S. (1984). Regulation of a hybrid gene by glucose and temperature in hamster fibroblasts. Science **226**, 187–190.

BOLE, D.G., HENDERSHOT, L.M., and KEARNEY, J.F. (1986). Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in nonsecreting and secreting hybridomas. J. Cell Biol. **102**, 1558–1566.

BREATHNACH, R., BENOIST, C., O'HARE, K., GANNON, F., and CHAMBON, P. (1978). Ovalbumin gene: Evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. Proc. Natl. Acad. Sci. USA **75**, 4853–4857.

CHANG, S.C., WOODEN, S.K., NAKAKI, T., KIM, Y.K., LIN, A.Y., KUNG, L., ATTENELLO, J.W., and LEE, A.S. (1987). Rat gene encoding the 78-kDa glucose-regulated protein GRP78: Its regulatory sequences and the effect of protein glycosylation and its expression. Proc. Natl. Acad. Sci. USA **84**, 680–684.

CHAPPELL, T.G., WELCH, W.J., SCHLOSSMAN, D.M., PALTER, K.B., SCHLESINGER, M.J., and ROTHMAN, J.E. (1986). Uncoating ATPase is a member of the 70 kilodalton family of stress protein. Cell **45**, 3–13.

GETHING, M.-J., McCAMMON, K., and SAMBROOK, J. (1986). Expression of wild-type and mutant forms of influenza hemagglutinin: The role of folding in intracellular transport. Cell **46**, 939–950.

HAYASHIDA, H., and MIYATA, T. (1983). Unusual evolutionary conservation and frequent DNA segment exchange in class I genes of the major histocompatibility complex. Proc. Natl. Acad. Sci. USA **80**, 2671–2675.

HENDERSHOT, L., BOLE, D., KOHLER, G., and KEARNEY, J.F. (1987). Assembly and secretion of heavy chains that do not associate posttranslationally with immunoglobulin heavy chain-binding protein. J. Cell Biol. **104**, 761–767.

HOLLIS, G.F., HIETER, P.A., McBRIDE, O.W., SWAN, D., and LEDER, P. (1982). Processed genes: A dispersed human immunoglobulin gene bearing evidence of RNA-type processing. Nature **296**, 321–325.

KARIN, M., and RICHARDS, R.I. (1982). Human metallothionein genes-primary structure of the metallothionein-II gene and a related processed gene. Nature **299**, 797–802.

KLEINSEK, D.A., BEATTIE, W.G., TSAI, M.-J., and O'MALLEY, B.W. (1986). Molecular cloning of a steroid-regulated 108 k heat shock protein gene from hen oviduct. Nucleic Acids Res. **14**, 10053–10069.

KYTE, J., and DOOLITTLE, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**, 105–132.

LAWN, R.M., FRITSCH, E.F., PARKER, R.C., BLAKE, G., and MANIATIS, T. (1978). The isolation and characterization of linked δ- and β-globin genes from a cloned library of human DNA. Cell **15**, 1157–1174.

LEE, A.S. (1981). The accumulation of three specific proteins related to glucose-regulated proteins in a temperature-sensitive hamster mutant cell line K12. J. Cell. Physiol. **106**, 119–125.

LEE, A.S. (1987). Coordinated regulation of a set of genes by glucose and calcium ionophores in mammalian cells. Trends Biochem. Sci. **12**, 20–23.

LEE, A.S., DELEGEANE, A., and SCHARFF, D. (1981). Highly conserved glucose-regulated protein in hamster and chicken cells: Preliminary characterization of its cDNA clone. Proc. Natl. Acad. Sci. USA **78**, 4922–4925.

LEE, A.S., DELEGEANE, A.M., BAKER, V., and CHOW, P.C. (1983). Transcriptional regulation of two genes specifically induced by glucose starvation in a hamster mutant fibroblast cell line. J. Biol. Chem. **258**, 597–603.

LEE, A.S., WELLS, S., KIM, K.S., and SCHEFFLER, I.E. (1986). Enhanced synthesis of the glucose/calcium-regulated proteins in a hamster cell mutant deficient in transfer of oligosaccharide core to polypeptides. J. Cell. Physiol. **129**, 277–282.

LIEBHABER, S.A., and KAN, Y.W. (1982). Different rates of mRNA translation balance the expression of the two human α-globin loci. J. Biol. Chem. **257**, 11852–11855.

LIN, A., and LEE, A.S. (1984). Induction of two genes by glucose starvation in hamster fibroblasts. Proc. Natl. Acad. Sci. USA **81**, 988–992.

LIN, A.Y., CHANGE, S.C., and LEE, A.S. (1986). A calcium ionophore-inducible cellular promoter is highly active and has enhancerlike properties. Mol. Cell. Biol. **6**, 1235–1243.

LINDQUIST, S. (1986). The heat-shock response. Annu. Rev. Biochem. **55**, 1151–1191.

MIYATA, T., YASUNAGA, T., and NISHIDA, T. (1980). Nucleotide sequence divergence and functional constraint in mRNA evolution. Proc. Natl. Acad. Sci. USA **77**, 7328–7332.

MUNRO, S., and PELHAM, H.R.B. (1986). An Hsp-70 like protein in the endoplasmic reticulum identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell **46**, 291–300.

MUNRO, S., and PELHAM, H.R.B. (1987). A C-terminal signal prevents secretion of luminal ER proteins. Cell **48**, 899–907.

PELHAM, H.R.B. (1986). Speculation on the functions of the major heat shock and glucose-regulated proteins. Cell **46**, 959–961.

PINHASI-KIMHI, O., MICHALOVITZ, D., BEN-ZEEV, A., and OREN, M. (1986). Specific interaction between the p53 cellular tumour antigen and major heat shock proteins. Nature **320**, 182–185.

RESENDEZ, E., JR., ATTENELLO, J.W., GRAFSKY, A., CHANG, C.S., and LEE, A.S. (1985). Calcium ionophore A23187 induces expression of glucose-regulated genes and their heterologous fusion genes. Mol. Cell. Biol. **5**, 1212–1219.

RESENDEZ, E., JR., TING, J., KIM, K.S., WOODEN, S.K., and LEE, A.S. (1986). Calcium ionophore A23187 as a regulator of gene expression in mammalian cells. J. Cell Biol. **103**, 2145–2152.

ROSCOE, D.H., READ, M., and ROBINSON, H. (1973). Isolation of temperature-sensitive mammalian cells by selective detachment. J. Cell. Physiol. **82**, 325–332.

SANGER, F., NICKLEN, S., and COULSON, A.K. (1977). DNA sequencing with chain-terminating inhibitor. Proc. Natl. Acad. Sci. USA **74**, 5463–5467.

SORGER, P.K., and PELHAM, H.R.B. (1987). Cloning and expression of a gene encoding hsc73, the major hsp70-like protein in unstressed rat cells. EMBO J. **6**, 993–998.

TING, J. (1988). "Biochemical and molecular analysis of two glucose-regulated proteins (GRPs)." Ph.D. dissertation (University of Southern California, Los Angeles).

TING, J., WOODEN, S.K., KRIZ, R., KELLEHER, K., KAUFMAN, R.J., and LEE, A.S. (1987). The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protein (GRP78) and its conservation between hamster and rat. Gene **55**, 147–152.

VAN BEVEREN, C., VAN STRAATEN, F., CURRAN, T., MULLER, R., and VERMA, I.M. (1983). Analysis of FBJ-MuSV provirus and c-*fos* (mouse) gene reveals that viral and cellular *fos* gene products have different carboxyl terminals. Cell **32**, 1241–1255.

VAN STRAATEN, F., MULLER, R., CURRAN, T., VAN BEVEREN, C., and VERMA, I.M. (1983). Complete nucleotide sequence of a human c-*onc* gene: Deduced amino acid sequence of the human c-*fos* protein. Proc. Natl. Acad. Sci. USA **80**, 3183–3187.

WILDE, C.D., CROWTHER, C.E., CRIPE, T.P., LEE, M.G., and COWAN, N.J. (1982). Evidence that a human β-tubulin pseudogene is derived from its corresponding mRNA. Nature **297**, 83–84.

WU, B., HUNT, C., and MORIMOTO, R. (1985). Structure and expression of the human gene encoding major heat shock protein HSP70. Mol. Cell. Biol. **5**, 330–341.

YAFFE, D., NUDEL, U., MAYER, Y., and NEUMAN, S. (1985). Highly conserved sequences in the 3′ untranslated region of mRNAs coding for homologous proteins in distantly related species. Nucleic Acids Res. **13**, 3723–3737.

ZALA, C.A., SALAS-PRATO, M., YAN, W.T., BANJO, B., and PERDUE, J.F. (1980). In cultured chick embryo fibroblasts the hexose transport components are not the 75,000 and 95,000 dalton polypeptide synthesized following glucose deprivation. Can. J. Biochem. **58**, 1179–1188.

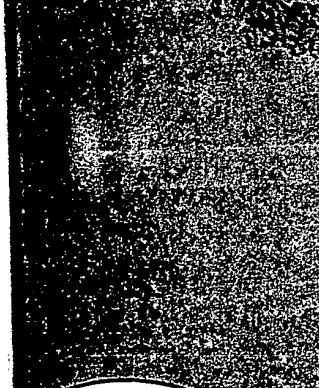
Address reprint requests to:
Dr. Amy S. Lee
Department of Biochemistry and
Norris Cancer Research Institute
University of Southern California
School of Medicine
1441 Eastlake Avenue
Los Angeles, CA 90033

Received for publication December 1, 1987, and in revised form January 25, 1988.

EXHIBIT B

VOLUME 8 ○ OCTOBER 1988 ○ NUMBER 10

Molecular and Cellular Biology



Published monthly by the American Society for Microbiology

CODEN: MCEBD4 8 (10) 3975–4585 (1988) ISSN: 0270-7306

MOLECULAR AND CELLULAR BIOLOGY, Oct. 1988, p. 4579-4584
0270-7306/88/104579-06$02.00/0

Vol. 8, No. 10

Identification of Highly Conserved Regulatory Domains and Protein-Binding Sites in the Promoters of the Rat and Human Genes Encoding the Stress-Inducible 78-Kilodalton Glucose-Regulated Protein

ELPIDIO RESENDEZ, JR., SCOTT K. WOODEN, AND AMY S. LEE*

Department of Biochemistry and the Norris Cancer Research Institute, University of Southern California School of Medicine, Los Angeles, California 90033

Received 23 March 1988/Accepted 24 June 1988

The gene encoding GRP78 has been shown to be constitutively expressed in many cell types and is inducible by the calcium ionophore A23187. To understand the regulation of GRP78 transcription, we analyzed the components that control its basal-level expression. By transfecting deletions into cells, we have identified a 54-nucleotide *cis*-acting regulatory element important for high basal-level expression and a contiguous 50-nucleotide element for both basal-level expression and A23187 induction. Using DNase footprinting assays with both rat and human GRP78 promoters, we demonstrated that the protein factors present in the HeLa cell nuclear extracts bind to the regulatory regions identified by the deletion studies. This domain contains a palindromic sequence and is highly conserved among GRP genes in *Caenorhabditis elegans*, chicks, rats, and humans.

A set of stress-inducible proteins known as the glucose-regulated proteins (GRPs) are constitutively expressed in mammalian cells. When the cells are deprived of glucose or treated with reagents that inhibit protein glycosylation, perturb intracellular calcium stores, or denature proteins, the synthesis of the GRPs is rapidly increased (11). The major GRP in mammalian cells has a molecular mass of about 78,000 daltons and is localized within the endoplasmic reticulum (ER) (18, 25a). This protein, generally referred to as GRP78, shares about 60% sequence homology with the 70,000-dalton heat-shock protein (HSP70) and is thought to have a function analogous to that of HSP70 but within the ER (18). By immunoprecipitation and comparisons of peptide maps and amino-terminal sequences, GRP78 was found to be identical to the immunoglobulin-binding protein referred to as BiP (1, 4, 15). In lymphoid and fibroblast cells, GRP78-BiP has been shown to bind to immature immunoglobulin and aberrant proteins (1, 18). Recently, the sequence of the 3,215-dalton steroidogenesis activator polypeptide isolated from a rat cell tumor has been reported (17). Strikingly, this 30-residue peptide, which has the ability to facilitate cholesterol side-chain cleavage to pregenolone, is identical, except for two residues, with the last 30 amino acids of the GRP78 carboxyl terminus as predicted from the DNA sequence of the rat GRP78 gene (15, 23).

The gene encoding GRP78 has recently been isolated from a human genomic library (22a). Both the rat and the human GRP78 genes are split into eight exons and contain a signal sequence which targets GRP78 into the ER (22a, 25a). The expression of the GRP genes is regulated at the transcriptional level (9, 10, 13, 19). Thus, when the cells are treated with potent inducers such as the calcium ionophore A23187, β-mercaptoethanol, or tunicamycin, the transcription of the GRP78 gene is increased 10- to 25-fold within 5 h (3, 9, 19). The GRP78 promoter is highly active, and a 291-nucleotide (nt) fragment from the promoter region functions as an enhancer when fused to other cellular promoters (12). This region of the GRP78 promoter is G+C rich and contains several CCAAT sequences and extensive arrays of tandem and inverted repeat motifs characteristic of other cellular and viral enhancers (6). In addition, this 291-nt fragment, when placed upstream of a heterologous gene, can confer inducibility by calcium ionophore (12).

GRP78 transcript levels in HeLa cells. To facilitate the search for cellular factors which interact with the DNA regulatory domain, we utilized the HeLa cell system, which can be grown in suspension to provide high yields of cellular extracts. We have shown previously that GRP78 transcripts can be detected in a variety of mammalian cell lines (20). Since the focus of this study was on the utilization of the human HeLa S3 cell line as a source of functional protein extracts, we first investigated the expression of the endogenous GRP78 gene in these cells. GRP78 mRNA was detectable in noninduced HeLa cells grown in suspension cultures (Fig. 1). When the cells were further treated with 0.5 or 2 μM of the calcium ionophore A23187 or were grown continuously in the culture medium without a medium change, the endogenous GRP78 transcript levels were further increased (Fig. 1). These results demonstrate that the GRP78 gene is constitutively expressed in HeLa cells to provide a basal level of GRP78 transcripts. Furthermore, A23187 is capable of eliciting a strong stress response in the HeLa cells. However, A23187 is relatively toxic to HeLa S3 cells compared with fibroblast cell lines from hamster, rat and mouse. The optimum conditions for the induction of the GRP78 gene in HeLa S3 cells are 2 μM A23187 and an incubation period not longer than 6 h. Higher concentrations of A23187 or longer incubation periods or both result in substantial cell death (unpublished results).

Transcriptional activities of rat GRP78 promoter-deletion mutants in human cells. It has been shown previously that a 291-nt SmaI-StuI fragment upstream of the TATA element of the rat GRP78 promoter is important for both high basal level and A23187 induction of the reporter gene after transfection

* Corresponding author.

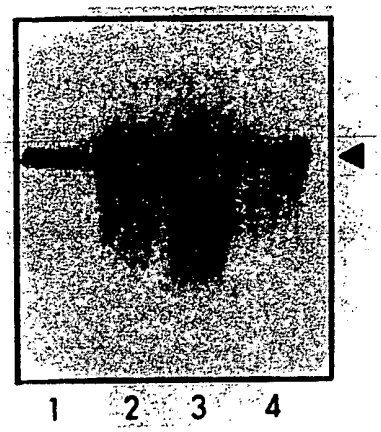


FIG. 1. Enhanced expression of GRP78 mRNA levels in HeLa cells. Cytoplasmic RNA was extracted from HeLa cells after incubation under the following conditions: control (lane 1), 0.5 μM A12387 (lane 2), 2 μM A23187 (lane 3), and medium starvation (lane 4). For control and calcium ionophore (A23187) treatment, cells were placed in fresh medium 8 h before cytoplasmic RNA extraction; for calcium ionophore treatment, cells were incubated for 5 h with the ionophore before RNA extraction; and for the medium starvation condition, the cells were kept in the original medium and incubated for an additional 24 h before RNA extraction. The cytoplasmic RNA samples (5 μg) were electrophoresed on a formaldehyde denaturing gel, trans-blotted onto a nitrocellulose filter, and hybridized with radiolabeled hamster GRP78 cDNA plasmid p3C5 (23a) as described previously (10, 20). The arrow indicates the position of the 2.7-kilobase GRP78 transcript.

into hamster fibroblast K12 cells (2, 12). The sequence of this region and the proximal promoter sequences are shown in Fig. 2A. To delineate further the functional domains within this region, a series of 5′-deletion mutants were constructed and their deletion endpoints were determined by DNA sequencing. To test the functionality of these deleted 5′-flanking sequences, they were fused upstream to the bacterial chloramphenicol acetyltransferase (CAT) transcriptional unit (Fig. 2B). These 5′-deletion plasmids were transfected into HeLa, HepG2, and K12 cells. The cells were transfected with 3 or 5 μg of cesium chloride gradient-purified plasmid DNA with 5 or 7 μg of high-molecular-weight carrier HeLa DNA. The transfection conditions and the assay of CAT activity were performed as described previously (16, 19, 24). The promoter strengths of the deletion mutants were determined by the CAT enzymatic activities under noninduced and A23187-induced conditions (Table 1). In these transient transfection experiments, we failed to detect CAT activities in the HeLa cells transfected with the CAT plasmids, probably due to the low transfection efficiency of the HeLa cells used. However, our results with the human HepG2 and hamster K12 cells consistently showed that between deletions −208 and −154, a fivefold reduction in the basal level of expression was observed, suggesting that this 54-nt domain contains important *cis*-acting regulatory elements for high basal-level expression. When the promoter is deleted to −130, the basal level was further reduced and the 2.5- to 8-fold induction by A23187 observed in pGRP78 (−154)CAT was reduced by one-half. In the deletion mutants pGRP78(−104)CAT and pGRP78(−85)CAT, only minimal activities were detected. While these activities were higher than the promoterless plasmid pSV0CAT, the basal level was further reduced and A23187-induced expression was eliminated. This implies that the 50-nt domain between deletions −154 and −104 contains important DNA elements for both basal level and A23187 induction.

TABLE 1. Relative CAT activities[a]

Deletion endpoint	HepG2		K12	
	Basal level	Level with A23187 (increase [fold])	Basal level	Level with A12387 (increase [fold])
−208	100	150 (1.5)	100	500 (5.0)
−154	20	50 (2.5)	20	160 (8.0)
−130	3	4 (1.3)	6	20 (3.5)
−104	2	2 (1.0)	2	3 (1.5)
−85	1.5	1.5 (1.0)	2	3 (1.5)
pSV0CAT	0.5	0.5 (1.0)		

[a] Level of CAT activity of pGRP78(−208)CAT under noninduced conditions was set as 100. Results represent averages of three sets of transfections for each cell line.

The HeLa extract contains protein factors that bind to the rat GRP78 regulatory domain. Using an in vitro transcription assay, we observed that the nuclear extracts from HeLa cells were able to transcribe the rat GRP78 promoter in a template- and extract-dependent manner. In addition, by gel mobility shift experiments we detected multiple complexes formed between HeLa nuclear extract and the regulatory region of the GRP78 promoter defined above (unpublished

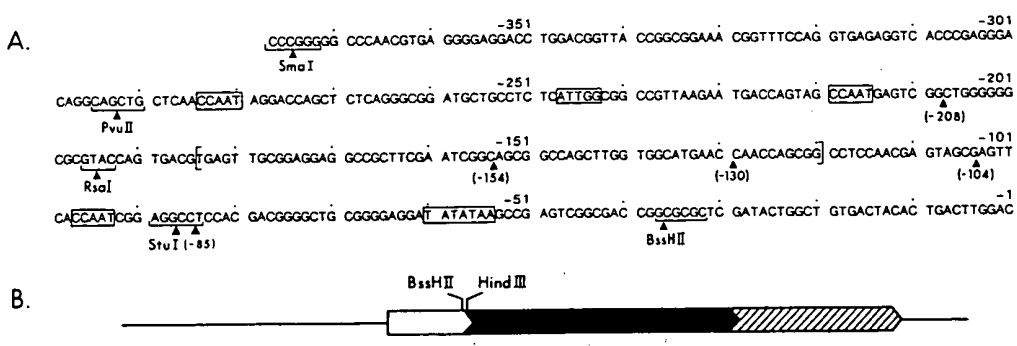


FIG. 2. Promoter deletions and activities of CAT constructs. (A) The sequence of the rat GRP78 promoter. Bases are numbered in reference to the major mRNA cap site of GRP78 (12). Several relevant restriction sites are noted. The deletion mutants were generated by BAL 31 digestion of pI10 (19). The deletion endpoints for each of the deletion plasmids used for the transfection are indicated. The CCAAT and TATA regions are boxed. The region protected from DNase I digestion described below is bracketed. (B) The organization of the deletion mutant plasmids. The GRP78-deleted promoter with an uniform 3′ deletion endpoint at the *Bss*HII site is fused to the *Hin*dIII site of pSV0CAT (19). Other regions of the plasmid include: the CAT gene (■), the SV40 polyadenylation site (▨), and sequences from pBR322 (—).

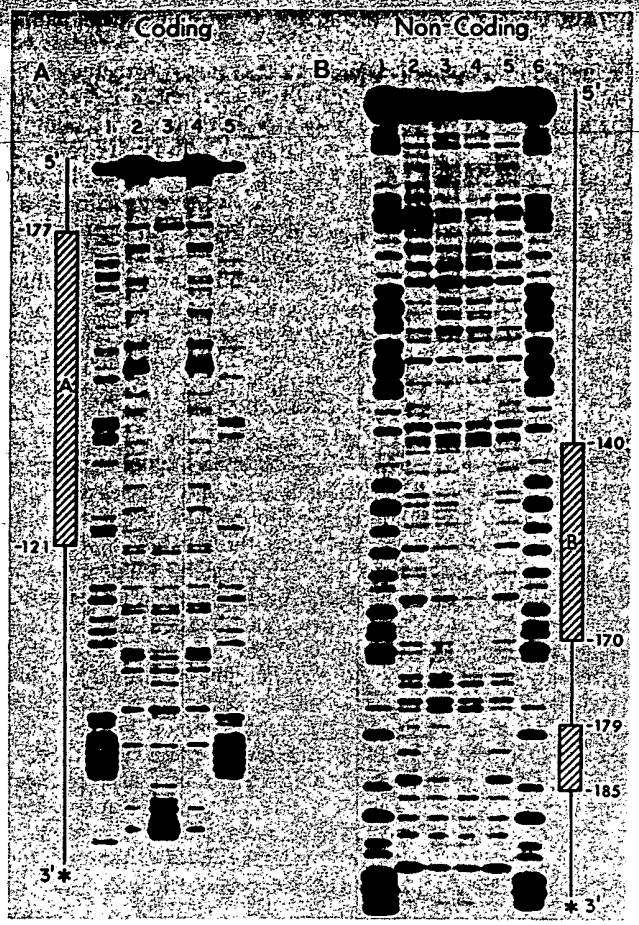


FIG. 3. Footprint analysis of the rat GRP78 promoter. A 159-nt EcoRI-HindIII fragment (spanning −195 to −88) was end labeled, cut with *Asp*718 to obtain the labeled coding strand or with *Sal*I to obtain the labeled noncoding strand, and gel purified. The HeLa nuclear protein extract was prepared as described previously (21). (A) Coding strand footprint. Lanes: 2 and 4, DNA probe incubated without protein extract; 3, DNase-treated probe with 44 μg of HeLa nuclear protein extract; 1 and 5, Maxam and Gilbert G sequencing reaction (14) of the labeled strand. (B) Noncoding strand footprint. Lanes: 2 and 5, DNase-treated probe without protein extract; 3 and 4, 44 and 66 μg, respectively, of HeLa extracts; 1 and 6, Maxam and Gilbert G sequencing reactions. Diagrams of the coding and noncoding regions are shown next to the gels. The footprint regions A and B are boxed. The numbers denote nucleotides correspond to the locations of the footprints relative to the transcription start site (+1) of the rat GRP78 gene.

results). To define the binding sites of these factors within the rat GRP78 promoter, the subfragment spanning −195 to −88 was end labeled, mixed with HeLa nuclear extracts, and subjected to DNase I digestion. Binding reaction mixtures (20 μl) contained 10 mM Tris-hydrochloride (pH 7.5), 50 mM NaCl, 10 mM β-mercaptoethanol, 1 mM EDTA, 4% glycerol, 2 mM $MgCl_2$, 1 μg poly(dI-dC), and 0.4 to 1 ng of end-labeled DNA. After 20 min of incubation at room temperature, 20 μl of 5 mM $CaCl_2$–1 mM EDTA was added; freshly prepared DNase I was added, and the mixture was incubated for 60 s at room temperature. To terminate the nuclease reaction, 50 μl of buffer (200 mM NaCl, 20 mM EDTA, 1% sodium dodecyl sulfate, 250 μg of tRNA per ml) was added. Samples were deproteinized by phenol-chloroform extraction, and the DNA was ethanol precipitated, suspended in formamide-dye mix, and electrophoresed on a denaturing 6% polyacrylamide–8M urea sequencing gel. The analysis was performed for both the coding and the noncoding strands of the DNA. The results revealed that a region spanning positions −121 to −177 (Fig. 3, A box) was protected on the coding strand. In the case of the noncoding strand, two regions spanning positions −140 to −170 (Fig. 3, B box) and −179 to −185 of the noncoding strand were

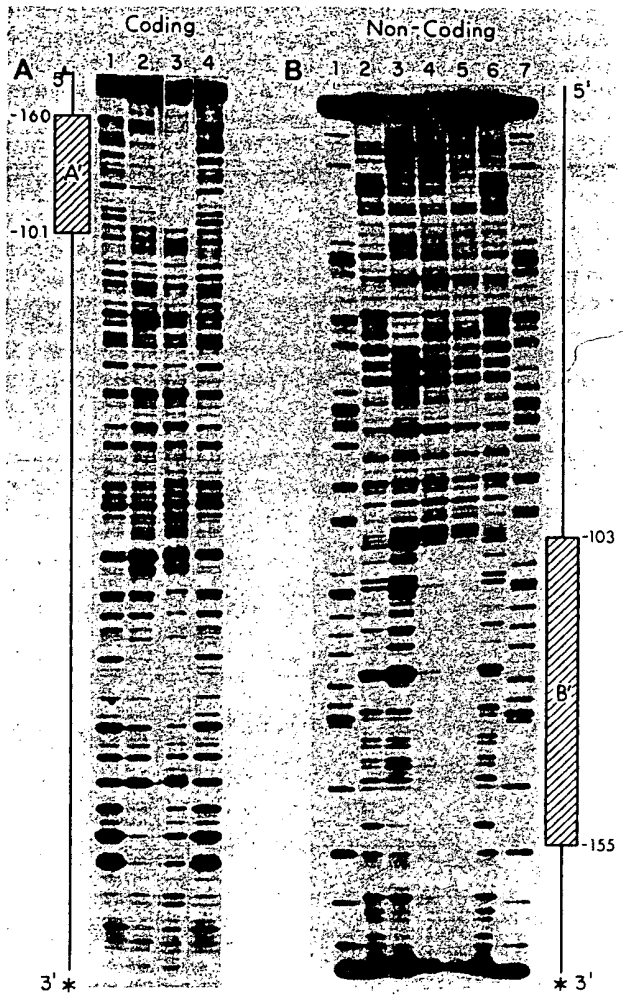


FIG. 4. Footprint analysis of the human GRP78 promoter. pUCH223 (−170, +53), which contains the 223-nt PvuII-NruI fragment from the human GRP78 promoter (22a) subcloned into the *Sma*I site of pUC8 (with the *Pvu*II site proximal to the *Eco*RI site of the pUC8 polylinker sequence), was linearized with *Hin*dIII, end labeled with T4 DNA polymerase, and cut with *Eco*RI to obtain the labeled coding strand. The labeled noncoding strand was obtained by labeling the *Eco*RI-*Hin*dIII insert of pUCH223 and recutting with *Sal*I. (A) Coding strand footprint. Lanes: 1 and 4, DNase-treated probe without nuclear extract; 2 and 3, DNase-treated probe with 44 and 66 μg, respectively, of HeLa nuclear protein. (B) Noncoding strand footprint. Lanes: 2 and 6, DNase-treated probe without nuclear extract; 3, 4, and 5, DNase-treated probe with 22, 44, and 66 μg, respectively, of HeLa nuclear protein; 1 and 7, Maxam and Gilbert G sequencing reaction of the labeled strand. Diagrams of the coding and noncoding regions are shown next to the gels. The boxes represent the footprint regions (A′ and B′). The numbers denote nucleotides and correspond to the locations of the footprints relative to the transcription start site (+1) of the rat GRP78 gene.

protected from DNase I digestion. These domains reside within the functional domains (−104 to −208) defined by in vivo transfection experiments described above.

Binding of protein factors to the human GRP78 promoter. We have previously shown that the rat and the human GRP78 promoter sequences are highly conserved (22a). The 170 nt proximal to the TATA sequence in the human GRP78 promoter have been identified as important for basal-level expression and inducibility by A23187 (22a). Sequences further upstream enhance the basal-level expression by twofold. Therefore, it is of interest to determine whether the human GRP78 promoter interacts with protein factors from HeLa cells in a similar pattern as the rat promoter. For this purpose, a human GRP78 promoter subfragment spanning from −170 to +53 of the promoter was end labeled at either its coding or noncoding strand, mixed with the HeLa extracts, and subjected to DNase I digestion as described above. The results indicate that regions spanning residues −101 to −160 are protected in the coding strand (Fig. 4, A′ box) and residues −103 to −155 are protected in the noncoding strand (Fig. 4, B′ box). The A′ and B′ sequences from the human GRP78 promoter are similar to the A and B sequences of the rat GRP78 promoter. Another region more proximal to the TATA element also exhibited some protection in both strands at high DNase I concentration. This region was not detected in the rat GRP 78 promoter shown in Fig. 3, since the rat DNA fragment used for the footprint analysis was shorter than the human fragment. In other experiments in which a longer rat promoter fragment was used, the same region was protected (T. Nakaki and A. S. Lee, unpublished results).

Comparison of rat and human GRP78 promoter regions that bind to cellular factors. The sequences of the rat and human GRP78 promoters binding to proteins are summarized in Fig. 5. Generally, the same region protected by DNase I digestion in the rat promoter is also protected in the human promoter. This is particularly evident in the case of the noncoding strand, where one end point of protection is identical and the other end is similar. The human promoter exhibits a footprint on its coding strand which directly overlaps that of its noncoding strand. However, the footprint (Fig. 4, A box) of the coding strand of the rat promoter is longer than the noncoding footprint (Fig. 4, B Box) by about 10 to 20 nt. These subtle differences may be related to the DNA sequence divergence of the two promoters. However, the sequences at the protection boundaries are identical for the coding strands of the two promoters. Most importantly, the protected domains coincide with the region shown to be critical for in vivo expression of the GRP78 gene. Within this region, a palindromic sequence and three pairs of short direct repeats are present.

The GRP78 regulatory domain is highly conserved. The above in vivo deletion and the footprint protection results both point to a small domain within the rat GRP78 promoter that interacts with proteins from HeLa nuclear extracts. The availability of sequence data from several GRP genes allowed us to compare their sequences for possible common consensus (Fig. 5). While the rat and human GRP78 genes are highly conserved over a 200-nt region of the two promoters (22a), the GRP78 promoter sequence from *Caenorhabditis elegans* is very divergent from the mammalian sequence (M. Heschl and D. Baillie, submitted for publication). However, there is one region within the *C. elegans* promoter which has 19 out of 23 nt matched with the rat sequence. As shown in Fig. 5, this region resides within the critical domain for high basal-level expression defined by our analysis. On

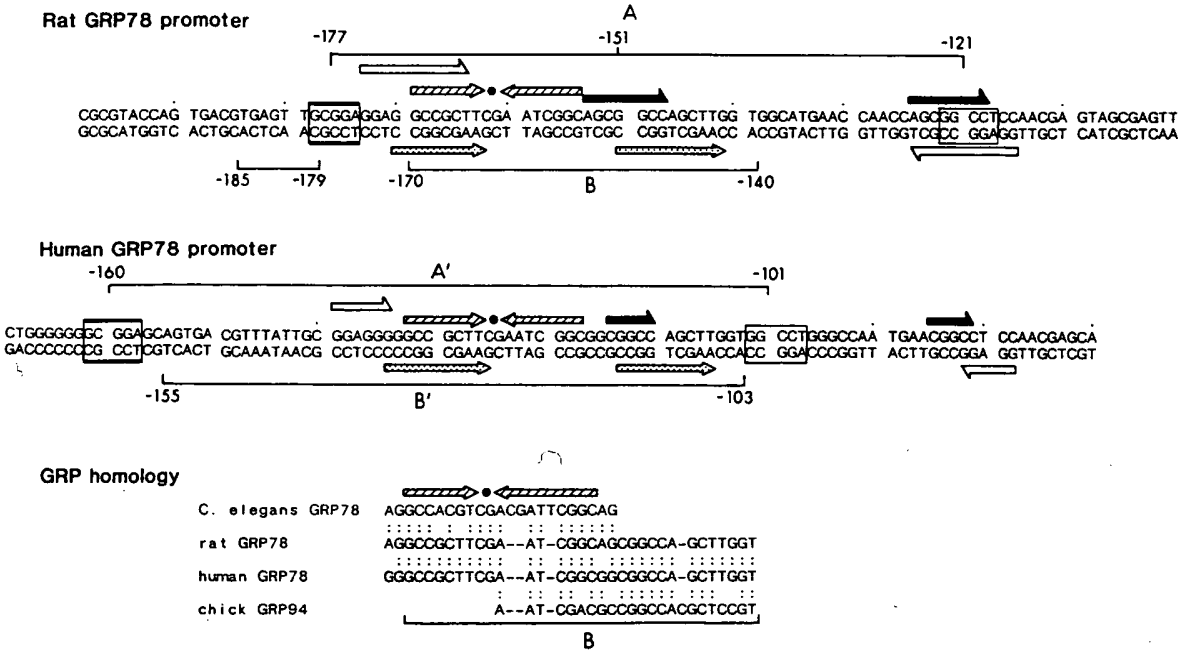


FIG. 5. Footprint sequence of the rat and human DNA and its conservation. The sequences that footprint on the coding strand of rat (A) and human (A′) GRP78 gene and on the noncoding strand of rat (B) and human (B′) GRP78 gene are bracketed. The palindromic sequence is indicated by a pair of hatched arrows, with the center of symmetry highlighted by a black dot. Other pairs of direct repeats are indicated by the other types of arrows. The homologous sequences between the rat and human promoter at the coding strand footprint boundaries are boxed. The highly conserved domain of the rat GRP78 promoter with the 5′-flanking regions of GRP genes of the other species is also shown. The location of the palindromic sequence and the footprint B domain shown above for the rat GRP78 promoter are indicated. Identical nucleotides are indicated by the vertical dots.

the other hand, GRP94, which is expressed at a lower basal level than GRP78 but is inducible by A23187, has 18 out of 23 nt matched with the rat and human GRP78 promoters within the region important for A23187 induction (22a). These two domains are also the same domain (Fig. 3, B box) detected by our footprint analysis. The strict conservation of these short domains within such phylogenetically diverse species as worms, chicks, rats, and humans strongly implies that they have a functional significance in the regulation of GRP gene expression.

The GRP78 gene system represents a unique model for the study of the regulation of genes encoding ER-localized proteins. Most interestingly, GRP78 transcriptional activity appears to correlate directly with the amount of ER activity or ER damage in a variety of cell types, bringing up the question of how the state of another organelle, such as the ER, can regulate a gene in the nucleus. Clearly, intermediate molecules must exist and traverse the membranes to communicate the signal to the GRP gene system. Since the ER membranes are associated with the perinuclear membranes, traffic between the two organelles can be envisioned.

By 5'-deletion analysis, we show here that a 100-nt sequence within this region is critical for both high basal and A23187-induced expression. This same DNA segment is highly conserved among the GRP genes and is protected by cellular factors in DNase protection assays. If the footprinted domain represents the core for the binding activities of cellular factors that regulate GRP78 gene expression, this region may represent binding sites for a novel class of transcriptional factors that have not yet been described. By sequence analysis, it is devoid of recognition sites for characterized transcriptional factors such as Sp1, CTF, AP1, AP2, and TFIID (5). Since this region contains a pair of inverted repeats which basically results in the same sequence independent of orientation, it may explain the orientation-independent characteristics of the DNA enhancer element contained within the GRP78 promoter.

Although GRP78 shares partial sequence homology with members of the 70,000-dalton heat shock protein family and may share the properties of binding to abnormal proteins under stressful conditions, the regulation of GRP78 expression is distinct from that of HSP70 in many ways. First, the transcriptional activation of the GRP78 gene is sensitive to the protein synthesis inhibitor cycloheximide (8, 20), whereas the induction of the heat-inducible HSP70 gene does not require de novo protein synthesis. In fact, it has been demonstrated that the heat shock regulatory factor can bind to the heat shock element in the presence of cycloheximide (26). In addition, the most potent inducers of GRP78 expression, such as calcium ionophores and β-mercaptoethanol, do not affect HSP70 expression (7, 9, 25). Therefore, whatever is the mechanism of induction of GRP78, it is likely to be different from that of the HSP70 gene and probably involves multiple steps (7, 11, 25). The identification of the *trans*-acting factors which regulate the expression of the GRP78 gene holds the key to uncovering the intermediate steps whereby a signal from the ER can be transmitted to the nucleus to enhance GRP78 transcription.

We thank Mark Heschl and David Baillie for communication of the *C. elegans* promoter sequence before publication and Walter Miller for alerting us to the similarity between SAP and GRP78.

E.R., Jr. was supported in part by a predoctoral fellowship (GM 9805-MC) from the National Institutes of Health, and S.W. was supported by NIH training grant T32 CA09569-01A1. This investigation was supported by Public Health Service Grant R37 CA 27607 awarded by the National Cancer Institute, Department of Human and Health Services. A. S. Lee is a recipient of a Faculty Research Award from the American Cancer Society.

The GenBank accession numbers are J03377 for the rat sequence and M19645 for the human sequence.

LITERATURE CITED

1. **Bole, D., L. Hendershot, and J. F. Kearney.** 1986. Post-translational association of immunoglobulin heavy chain binding protein with nascent heavy chains in non-secreting and secreting hybridomas. J. Cell Biol. **102**:1558–1566.
2. **Chang, S. C., S. K. Wooden, T. Nakaki, Y. K. Kim, A.-Y. Lin, L. Kung, J. W. Attenello, and A. S. Lee.** 1987. Rat gene encoding the 78 kDa glucose-regulated protein GRP78: its regulatory sequences and the effect of protein glycosylation on its expression. Proc. Natl. Acad. Sci. USA **84**:680–684.
3. **Drummond, I. S., A. S. Lee, E. Resendez, Jr., and R. A. Steinhardt.** 1987. Depletion of intracellular calcium stores by calcium ionophores A23187 induces the gene for the glucose-regulated proteins in hamster fibroblasts. J. Biol. Chem. **262**: 12801–12805.
4. **Hendershot, L., J. Ting, and A. Lee.** 1988. Identity of the immunoglobulin heavy-chain-binding protein with the glucose-regulated protein and the role of posttranscriptional modifications in its binding function. Mol. Cell. Biol. **8**:4250–4256.
5. **Jones, N. C., P. W. J. Rigby, and E. B. Ziff.** 1988. Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. Genes Dev. **2**:267–281.
6. **Khoury, G., and P. Gruss.** 1983. Enhancer elements. Cell **33**: 313–314.
7. **Kim, K. S., and A. S. Lee.** 1986. The effects of extracellular Ca^{++} and temperature on the induction of the heat shock and glucose-regulated protein in hamster fibroblast. Biochem. Biophys. Res. Comm. **140**:881–887.
8. **Kim, Y. K., K. S. Kim, and A. S. Lee.** 1987. Regulation of the glucose-regulated protein genes by β-mercaptoethanol requires de novo protein synthesis and correlates with inhibition of protein glycosylation. J. Cell Physiol. **133**:553–559.
9. **Kim, Y. K., and A. S. Lee.** 1987. Transcriptional activation of the glucose-regulated protein genes and their heterologous fusion by β-mercaptoethanol. Mol. Cell. Biol. **7**:2974–2976.
10. **Lee, A. S., A. M. Delegeane, V. Baker, and P. C. Chow.** 1983. Transcriptional regulation of two genes specifically induced by glucose starvation in a hamster mutant fibroblast cell line. J. Biol. Chem. **258**:597–603.
11. **Lee, A. S.** 1987. Coordinated regulation of a set of genes of glucose and calcium-ionophore in mammalian cells. Trends Biochem. Sci. **12**:20–23.
12. **Lin, A. Y., S. C. Chang, and A. S. Lee.** 1986. A calcium ionophore-inducible cellular promoter is highly active and has enhancer-like properties. Mol. Cell. Biol. **6**:1235–1243.
13. **Lin, A. Y., and A. S. Lee.** 1984. Induction of two genes by glucose starvation in hamster fibroblasts. Proc. Natl. Acad. Sci. USA **81**:988–992.
14. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
15. **Munro, S., and H. R. B. Pelham.** 1986. An hsp-70-like protein in the ER: identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell **46**:291–300.
16. **Parker, B. A., and G. R. Stark.** 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. J. Virol. **31**:360–369.
17. **Pedersen, R. C., and A. C. Brownie.** 1987. Steroidogenesis-activator for polypeptide isolated from a rat Leydig cell tumor. Science **236**:188–190.
18. **Pelham, H. R. B.** 1986. Speculations on the functions of the major heat shock and glucose-regulated proteins. Cell **46**:959–961.
19. **Resendez, E., Jr., J. W. Attenello, A. Grafsky, C. S. Chang, and A. S. Lee.** 1985. Calcium ionophore A23187 induces expression of glucose-regulated genes and their heterologous fusion genes. Mol. Cell. Biol. **5**:1212–1219.

20. **Resendez, E., Jr., J. Ting, K. S. Kim, S. K. Wooden, and A. S. Lee.** 1986. Calcium ionophore A23187 as a regulator of gene expression in mammalian cells. J. Cell. Biol. **103**:2145–2152.
21. **Shapiro, D. J., P. A. Sharp, W. W. Wahli, and M. J. Keller.** 1988. A high efficiency HeLa cell nuclear transcription extract. DNA 7:47–55.
22. **Sorger, P. K., and H. R. B. Pelham.** 1987. Cloning and expression of a gene encoding hsc73, the major hsp70-like protein in unstressed rat cells. EMBO J. **6**:993–998.
22a. **Ting, J., and A. S. Lee.** 1988. Human gene encoding the 78,000 dalton glucose-regulated protein and its pseudogene: structure, conservation, and regulation. DNA 7:275–286.
23. **Ting, J., S. K. Wooden, R. Kritz, K. Kelleher, R. J. Kaufman, and A. S. Lee.** 1987. The nucleotide sequence encoding the hamster 78-dDa glucose-regulated protein (GRP78) and its conservation between the hamster and rat. Gene **55**:147–152.
24. **Wang, M.-L., and A. S. Lee.** 1983. Polymerization of vector DNA after transfection into hamster fibroblast cells. Biochem. Biophys. Res. Commun. **110**:593–601.
25. **Watowich, S. S., and R. I. Morimoto.** 1988. Complex regulation of heat shock- and glucose-responsive genes in human cells. Mol. Cell. Biol. **8**:393–405.
25a. **Wooden, S. K., R. P. Kapur, and A. S. Lee.** 1988. The organization of the rat GRP78 gene and A12387 induced expression of fusion gene products targeted intracellularly. Exp. Cell Res. **178**:84–92.
26. **Zimarino, V., and C. Wu.** 1987. Induction of sequence-specific binding of *Drosophila* heat shock activator protein without protein synthesis. Nature (London) **327**:727–730.

EXHIBIT C

8H

MCB

MOLECULAR AND CELLULAR BIOLOGY

PUBLISHED
MONTHLY BY
THE AMERICAN
SOCIETY FOR
MICROBIOLOGY

MOLECULAR AND CELLULAR BIOLOGY, Aug. 1994, p. 5533–5546 Vol. 14, No. 8
0270-7306/94/$04.00+0

Stress Induction of the Mammalian GRP78/BiP Protein Gene: In Vivo Genomic Footprinting and Identification of p70CORE from Human Nuclear Extract as a DNA-Binding Component Specific to the Stress Regulatory Element

WILFRED W. LI,[1] LEA SISTONEN,[2]† RICHARD I. MORIMOTO,[2] AND AMY S. LEE[1]*

Department of Biochemistry and Norris Cancer Research Institute, University of Southern California School of Medicine, Los Angeles, California 90033,[1] and Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208[2]

Received 6 January 1994/Returned for modification 17 March 1994/Accepted 11 May 1994

GRP78, also known as BiP, is one of the better-characterized molecular chaperones. It has been implicated in protein folding and also calcium sequestration in the endoplasmic reticulum. When the cells are subjected to endoplasmic reticulum stress, in particular the depletion of stored calcium and/or the accumulation of abnormal proteins, the rate of transcription of *grp78* is enhanced. Previous studies have shown that the core region of the rat *grp78* promoter (−170 to −135), which is 95% conserved with the human *grp78* core (−133 to −98), is one of the key regulatory elements. Using ligation-mediated PCR, we have found that there are specific changes in factor occupancy after stress induction and the major changes occur within a cluster of bases located in the 3′ half of the *grp* core, whereas other regulatory elements are constitutively occupied. This inducible binding to the 3′ half of the human *grp78* core region is observed under diverse stress signals, suggesting a common mechanism for the *grp* stress response. Nonetheless, the lack of constitutive in vivo protection at this region is not due to the absence of a binding factor in nuclear extracts. Using in vitro gel mobility shift assays, we detected a constitutive binding activity which exhibits specificity and affinity to the stress-inducible region. Through sodium dodecyl sulfate-polyacrylamide gel electrophoresis size fractionation and renaturation analysis, the activity is found in polypeptides with molecular sizes of 65 to 75 kDa. After a three-step purification scheme including core affinity column chromatography, we purified p70CORE, which is about 70 kDa in its monomeric form. The purified p70CORE is sufficient to form a complex specific to the stress-inducible region.

The 78-kDa glucose-regulated protein (GRP78), also known as the immunoglobulin heavy-chain-binding protein (BiP), is a major cellular protein localized in the lumen of the endoplasmic reticulum (ER). It is a member of the 70-kDa heat shock protein (HSP70) protein family (46) and, in an ATP-dependent manner, assists in the translocation, folding, and assembly of oligomeric proteins (10, 17). In the capacity of a molecular chaperone protein, GRP78 binds to a wide repertoire of proteins traversing through the ER. Another function of GRP78 may involve its calcium-binding capabilities. It is postulated that GRP78, as well as other ER-localized GRPs, can bind and regulate the mobilization of sequestered calcium in the ER compartment, thus playing a key role in the calcium-dependent protein folding and transport processes (36, 41).

The *grp78* gene system provides an interesting model for studying the transcriptional activation of ER protein genes (16, 17). As a gene encoding a ubiquitous ER protein performing normal physiological functions, *grp78* is transcribed at a basal level constitutively in many different tissues and cell types (32). Its rate of transcription, however, is induced when the ER compartment is subjected to stress. Inducers for *grp78* include tunicamycin, which blocks glycosylation, denaturing reagents which disrupt protein disulfide bond formation, oxidative stress, the accumulation of malfolded or abnormal proteins in the ER (11, 15–17, 47), and continuous heat shock (43). Brefeldin A, which inhibits protein transport from the ER to the Golgi apparatus and causes Golgi resorption, activates *grp78* transcription (23). In mammalian cells, one of the most potent inducers for *grp78* transcription is the calcium ionophore A23187 (21, 32) and thapsigargin (19, 30, 44), a nonphorbol ester tumor promoter that inhibits specifically the endoplasmic reticulum Ca^{2+}-ATPase; both reagents enhance the transcription of *grp78* by about 20-fold through the depletion of Ca^{2+} stores from the ER (8). The efflux of Ca^{2+} from the ER store has also been known to interfere with protein folding (24).

The mammalian *grp78* promoter contains multiple elements which are functionally redundant (45). While the distal elements contribute mainly to the basal-level expression of the promoter, the proximal region contains the regulatory elements for stress inducibility (5). Using an extensive series of 5′ deletion, linker-scanning, and internal deletion mutants of the rat *grp78* promoter, we established that a region spanning −170 and −134 is important for basal-level expression and induction by A23187, thapsigargin, malfolded protein, and brefeldin (19, 23, 47). This region has been referred to as the *grp* core element because its sequence is conserved among *grp* promoters from many species, including *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, chickens, rats, and humans (22, 28, 34, 39). Especially between human and rat *grp78* core elements, the nucleotide sequence is 95% conserved. By using nuclear extracts from HeLa cells, this region is footprinted

* Corresponding author. Mailing address: USC/Norris Comprehensive Cancer Center, 1441 Eastlake Ave., P.O. Box 33800, Los Angeles, CA 90033-0800. Phone: (213) 224-6512. Fax: (213) 224-6417.
† Present address: Center for Biotechnology, 20521 Turku, Finland.

(34), and synthetic oligomers of the core elements can compete for nuclear factors binding to *grp* promoters in vitro and in vivo (5, 22). Stable integration and amplification of the core element in CHO cells coordinately downregulate the endogenous *grp* gene expression (20). Further, we demonstrated that synthetic oligomers of the *grp* core element are capable of conferring partial stress inducibility to a heterologous promoter (19).

In this study, we analyzed in vivo footprinting of the human *grp78* promoter and determined whether in vivo factor binding correlates with transcriptional induction of *grp78* by various forms of stress. We observed that there are specific changes in factor occupancy after stress induction and that the major changes occur within a cluster of bases located in the 3′ half of the *grp* core. Nonetheless, by gel mobility shift assays, factors capable of binding to the regulatory elements are present in cells prior to stress induction, irrespective of in vivo protection.

With the importance of the *grp* core established by genomic footprinting and transactivation of heterologous promoter in vivo, we seek to purify and characterize the nuclear proteins which interact with this site. Here we describe that in HeLa nuclear extracts, multiple factors bind to the core element. However, among the complexes formed with the *grp* core, a minor complex, termed II′, exhibits specific binding to the 3′-half region, which is required for stress induction of *grp78* (22) and is the site where factor occupancy changes are observed after stress induction in vivo. Complex II′ binds core DNA with high affinity and exhibits rapid association and dissociation kinetics. Previously, we identified several proteins from HeLa nuclear extracts with molecular sizes of 220, 110, 90/92, 70, and 55 kDa which were specifically cross-linked by UV light treatment to the *grp* core element (22). Here, by size fractionating the HeLa nuclear extract on sodium dodecyl sulfate (SDS)-polyacrylamide gels followed by renaturation, we observed that complex II′ is able to be reconstituted from the renatured proteins in the molecular size range of 65 to 75 kDa. Using a combination of chromatography steps, we purified a binding factor (p70CORE) from other binding factors, including the abundant Ku autoantigen, which exhibits high affinity for the core element. p70CORE, the first human nuclear factor identified to interact with the *grp78* core element, consists of protein species of 70 kDa in its monomeric form and binds specifically to the region required for A23187 inducibility and other kinds of stress induction of *grp78*.

MATERIALS AND METHODS

Cell lines and culture conditions. HeLa cells were grown in Dulbecco's modified Eagle's medium with 5 to 10% bovine calf serum. Cells were heat shocked at 42°C or treated with the calcium ionophore A23187 at a final concentration of 7 μM or with tunicamycin (0.5 μg/ml) for indicated periods of time. The hamster K12 cell line has been previously described (19, 33).

Transcriptional run-on analysis. Run-on transcription reactions were performed with isolated nuclei (approximately 7×10^6 nuclei per reaction) in the presence of 100 μCi of [α-^{32}P]UTP (3,000 Ci/mmol; Amersham) as previously described (3). Radioactive RNA was hybridized to nitrocellulose filters on which the following plasmids had been immobilized: pGEM2 (vector; Promega), pH 2.3 (human *hsp70* [48]), pUCHS801 (human *hsp90/89*α [12]), pHG23.1.2 (human *grp78* [42]), pHA7.6 (human *hsc70*), and pGAPDH (rat glyceraldehyde-3-phosphate dehydrogenase [9]). The hybridization and washing conditions were as described previously (38).

Genomic footprinting. For in vivo footprinting of the human *grp78* promoter, 2×10^7 cells were harvested and treated for 5 min with 0.2% dimethyl sulfate (DMS) at room temperature. Genomic DNA was isolated, digested with *Eco*RI, and cleaved with piperidine (4). The methylation pattern of naked DNA was obtained by using deproteinized DNA which was treated with DMS in vitro. Genomic footprinting analysis was performed by using a ligation-mediated PCR method (29). For the ligation-mediated PCR of the coding strand, the first genomic primer was a 19-mer, 5′ TGTCTGTGCTGTCTTGGCC 3′, which was complementary to the +12 to +30 region of the coding strand. The second genomic primer was a 25-mer, 5′ TTATATACCCTCCCCCAGCCCCGTC 3′, complementary to the −19 to −43 region of the coding strand. The third genomic primer used for end labeling was a 29-mer which overlapped and extended 3′ of the second primer and had the sequence 5′ TTATATACCCTCCCCCAGCCCCGTCGTGG 3′. The first genomic primer used for footprinting of the noncoding strand was a 19-mer, 5′ GGTCAGAAGTCGCAG GAGA 3′, complementary to the −283 to −265 region of the noncoding strand. The second genomic primer was a 25-mer, 5′ CTGAACCAATGGGACCAGCGGATGG 3′, complementary to the noncoding strand at nucleotides −252 to −228. The primer used for end labeling was a 28-mer which was complementary to the noncoding strand at nucleotides −252 to −225 and thus had the same sequence as the second primer with three additional bases, GGC, at the 3′ end. The sequence of the common linker was as described previously (29).

Preparation of nuclear extracts. The HeLa cells were cultured to a density of about 5×10^5 cells per ml. Exponentially growing cells were harvested to avoid growth density-related stress. For the A23187-induced nuclear extract, the HeLa cells were treated with 2 μM A23187 for 5 h prior to harvest. Nuclear extracts were prepared in parallel for the treated and nontreated cells as described previously (37), with the following modifications. A protease inhibitor cocktail was added to all buffers at the following final concentrations: phenymethylsulfonyl fluoride, 1 mM; benzamidine, 0.5 mM; and pepstatin, leupeptin, and aprotinin, 1 μg/ml. During cell lysis, the resuspended cell pellet was Dounce homogenized 8 to 20 times with tight pestle A (Wheaton Dounce tissue grinders), and the degree of lysis was monitored with a phase-contrast microscope. The nuclei were obtained by centrifugation with a Sorvall SS-35 rotor at 13,200 rpm for 30 s. Following ammonium sulfate precipitation, the nuclear extract was subjected to ultracentrifugation with a VTi70.1 rotor at 30,000 rpm for 20 min. After the final dialysis step, the extracts were quick-frozen in liquid nitrogen and stored in aliquots at −70°C.

Gel shift assays. The synthetic oligonucleotides were purified on 12% denaturing polyacrylamide gels, and equal amounts of the complementary strands were reannealed and labeled with Klenow enzyme in the presence of [α-^{32}P]dCTP (35). The probe was purified from free nucleotides by Microcon 10 columns (Amicon, Beverly, Mass.). Each gel shift reaction was carried out in 20-μl volume in the presence of 1× binding buffer {10 mM Tris (pH 7.5), 2 mM $MgCl_2$, 1 mM EGTA [ethylene glycol-bis (β-aminoethyl ether)-*N*,*N*,*N*′,*N*′,-tetraacetic acid], 80 mM NaCl, 0.25 g of bovine serum albumin (BSA) per ml}, with the few exceptions indicated in the figure legends. In regular binding studies, 0.1 μg of either control or induced HeLa nuclear extract, 1 ng of probe (specific activity, 10^7 to 10^8 cpm/μg), and 10 ng of sonicated salmon sperm DNA or 20 ng of sonicated poly(dI-dC) (average size, 200 bp) as a nonspecific competitor were added. However, no nonspecific DNA was used when DNA affinity-purified column fractions were assayed. In competition studies, nuclear extract and nonspecific DNA were added first, and then a molar excess of competitor oligonucleotides was added together with the

probe to the reaction mixture. In antibody shift studies, the nuclear extract was preincubated with antibodies for 3 to 5 min before the addition of probe. The complete reaction mixtures were incubated on ice for 20 to 30 min except where indicated. Electrophoresis was carried out in a Protein II xi apparatus (Bio-Rad, Richmond, Calif.) at either 150 V for 4.5 h or 300 V for 2.5 h. During electrophoresis, the temperature was maintained at 4 to 16°C through a recirculating ice water bath. The gel was dried, and the autoradiogram was exposed for 10 to 18 h at −70°C except where indicated.

Renaturation of proteins recovered from SDS-polyacrylamide gels. The procedure used was as described previously (2, 26), with the following modifications. Samples containing 200 μg of nuclear extracts were heated at 65°C for 3 min in SDS sample buffer (10 mM Tris [pH 6.8], 2% SDS, 5% glycerol, 2% β-mercaptoethanol, 0.1% bromophenol blue) and size fractionated by SDS-polyacrylamide gel electrophoresis (PAGE). Electrophoresis was carried out until the bromophenol blue dye ran out, in Protein II xi (16- by 20-cm plates) cooled to 4°C as described above. High-molecular-weight size markers (GIBCO-BRL) were run along with the samples for the calibration of the protein sizes in the sample lane, which was cut into 16 1- by 1-cm slices, crushed in Eppendorf tubes with a micro-tissue grinder, and soaked and rocked for 18 to 24 h in 500 μl of elution buffer (50 mM Tris [pH 7.9], 1 mM dithiothreitol, 0.2 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 0.1 mg of BSA per ml, 2.5% glycerol, 0.1% SDS) (26) at 4°C. The eluate was collected through centrifugation through 0.45-μm Z-spin columns (Gelman Sciences) at 14,000 × *g* for 5 min. The filtrate (250 μl) was acetone precipitated, washed with methanol, and air dried. Renaturation was carried out in renaturing buffer (20 mM Tris [pH 7.6], 10 mM KCl, 2 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride) at 4°C with rocking for 18 to 24 h. To store the renatured fractions, 50% glycerol in renaturing buffer was added to each tube to a final glycerol concentration of 25%. The samples were then quick-frozen in liquid nitrogen and stored at −70°C. To renature the affinity-purified nuclear extract, 100 μl of the affinity-purified fraction was used in a procedure identical to that described above.

Chromatographic fractionations of HeLa nuclear extracts. **(i) BioRex 70 chromatography.** All buffers in this and subsequent chromatography steps contained the protease inhibitors at concentrations specified above in the preparation of the extracts. All chromatographic steps were carried out at 4°C. The BioRex 70 resin (mesh size, 100/200) was preequilibrated in buffer D (20 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES; pH 7.9], 20% glycerol, 0.2 mM EDTA, 0.2 mM EGTA, 2 mM dithiothreitol) as recommended by the manufacturer (Bio-Rad). Fifty milligrams of HeLa nuclear extract prepared from A23187-treated cells was cleared by centrifugation at 10,000 × *g* for 10 min at 4°C. The proteins were then loaded onto a 12-ml column at 0.1 M KCl, incubated for 15 min, and then eluted with a KCl step gradient (0.2, 0.3, 0.4, 0.6, 0.8, and 1.0 M) in buffer D at 15 ml/h. The 0.4 M fractions containing II′ activity were pooled and dialyzed into buffer D with 0.1 M KCl. The concentration of the dialyzed proteins was determined by the Bio-Rad assay, with BSA as a standard. The protein profile was visualized by Coomassie blue staining of SDS-polyacrylamide gels.

(ii) DNA-cellulose chromatography. Double-stranded calf thymus DNA cellulose was obtained from Sigma. The material was preequilibrated in buffer D with 0.1 M KCl. The 0.4 M BioRex dialysate (2.1 mg) was applied onto a 9-ml column at a flow rate of 3 to 4 ml/h. The column was eluted with a step KCl gradient (0.2, 0.3, 0.4, 0.6, and 1.0 M) in buffer D. The 0.3 M fractions containing II′ activity were pooled. The protein concentration was determined by the Bio-Rad assay, and protein profiles were visualized by silver staining.

(iii) *grp* core affinity chromatography. The *grp78* core synthetic oligonucleotides were gel purified as described above; the complementary strands were reannealed and ligated as described previously (13). The oligomers (1 mg) were coupled to 2 g of cyanogen bromide-activated Sepharose 4B (Pharmacia, Piscataway, N.J.). The coupling of the oligomers to the resin and termination were performed as described previously (13, 49), and the resin was stored in the column storage buffer (10 mM Tris [pH 7.6], 1 mM EDTA, 0.3 M NaCl, 0.02% NaN_3) at 4°C. The 0.3 M fractions containing the II′ binding activity from the calf thymus DNA-cellulose column were pooled, dialyzed into buffer D with 0.1 M KCl, incubated with 2 ng of sonicated poly(dI-dC) per μl of extract, and applied onto a 1-ml (bed volume) affinity column. The column was eluted with a step gradient of KCl (0.1, 0.2, 0.4, 0.6, and 1.0 M) in buffer D containing 3 mM *n*-octyl glucoside (40) at a flow rate of 9 ml/h. The II′ binding activity was recovered in 0.4 M KCl eluate. The 0.4 M eluate was dialyzed against buffer D containing 0.1 M KCl and purified further by an additional cycle of core affinity chromatography. The purified fractions were stored in aliquots at −80°C. The protein profiles of the purified fractions were determined by silver staining, and the concentration was estimated by comparison against molecular weight markers.

RESULTS

Identification of constitutively associated transcription factors on the *grp78* promoter. In vivo genomic footprint analysis was used to identify which DNA sequences in the *grp78* promoter were constitutively associated with transcription factors. To accomplish this, we compared the methylation pattern of DNA isolated from non-heat-shocked cells with the pattern of deproteinized DNA methylated in vitro, i.e., naked DNA (Fig. 1). Protection against DMS reactivity mapped to guanine residues at positions −226, −224, −223, −207, −166 to −162, −160, −154, −152, −149, −125, and −89 on the top strand and at positions −182, −181, −170, −161, −156, −150, −126, −94, −93, −61, −60, −47, and −46 on the bottom strand. In addition, certain sites exhibited increased DMS reactivity; these included G-229, G-206, G-159, G-157, and A-87 on the top strand and G-53 and G-63 on the bottom strand. These alterations localize to the Sp1, CCAAT, and CREB-like sites present in the human *grp78* promoter and support earlier studies on the role of these sites for basal expression of *grp78* (47).

Transcriptional activation of *grp78* gene is accompanied by changes in protein-DNA interaction. We next compared the *grp78* promoter sequences in cells treated with A23187 for 3 h, treated with tunicamycin for 6 h, or exposed to a 42°C heat shock for 4 h. As shown in Fig. 2A, transcription of the *grp78* gene as measured by nuclear run-on analysis was induced. After exposure of the cells to A23187, tunicamycin, or heat shock, parallel cell samples were treated with DMS in vivo and the methylation pattern of genomic DNA was analyzed by genomic footprinting. The most prominent inducible alterations in DMS reactivity patterns in the *grp78* promoter can be detected at G-114, G-113, G-109, and G-75 on the coding strand and at G-112, G-111, and G-108 on the noncoding strand (Fig. 2B). Minor changes at G-104 and G-105 on the coding strand were also observed. The majority of changes in in vivo occupancy occur in a region of nucleotides −114 to −104, suggesting that the transcriptional induction of *grp78* is accompanied by factor binding to the sequence or that there

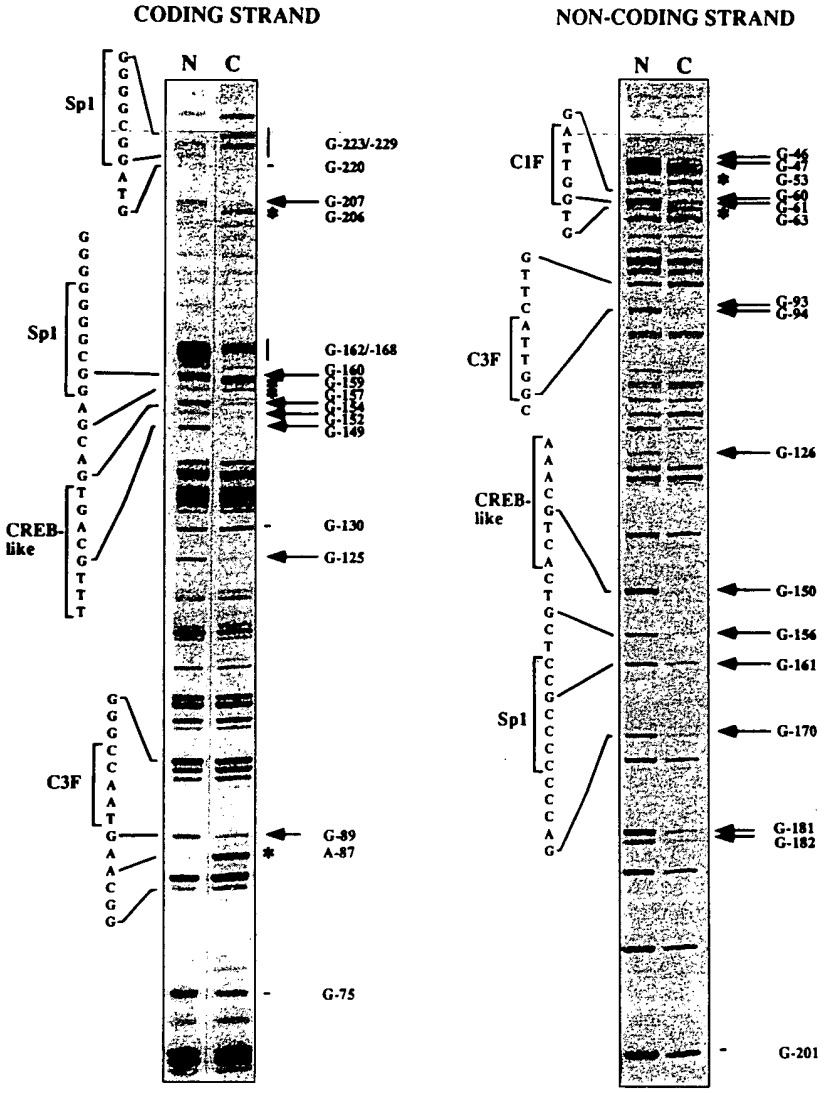


FIG. 1. Constitutive protein-DNA interactions within the human *grp78* promoter. Methylation patterns of the guanine (G) residues were analyzed by in vivo genomic footprinting and are presented for the coding and noncoding strands. Genomic DNA was isolated from control (C; no heat shock, A23187, or tunicamycin treatment) HeLa cells that were treated with DMS in vivo. Lane N contained naked (deproteinized) DNA that was treated with DMS in vitro. Long arrows indicate the G residues that are prominently protected from methylation, short arrows indicate partial protection, and asterisks indicate the G residues that are hypersensitive to methylation. The putative factor (C1F, C3F, CREB-like, and Sp1) binding sites are bracketed.

are changes in protein conformation occupying this site. The kinetics of inducible factor interactions correlated closely with the transcriptional activation and attenuation of the grp78 gene: the DMS methylation pattern in DNA isolated from cells that have recovered from the A23187 or tunicamycin treatment resembles the pattern seen in control cells, as does the level of *grp78* transcription (Fig. 2B, lanes R). In contrast to the transient nature of inducible factor interactions in the *grp78* promoter, the basal interactions located in the GC box, CCAAT box, and CREB-like elements were not affected by treatment with A23187 or tunicamycin. A summary of in vivo sites of factor interaction detected either in control or the induced state is shown in Fig. 3.

Binding of complex II′ to the core region important for A23187 induction. The cluster of changes in factor occupancy after stress occurred most prominently within the *grp* core, in particular the 3′ half of the core. The *grp* core element is highly conserved among *grp* promoters. A comparison of the human and rat *grp* core regions is shown in Fig. 4A. Within the 35 bp, there are only two base pair changes. The bases which showed changes in occupancy after stress are perfectly conserved between the two species. The rat core element has been further dissected into two functional domains (22). The 5′-half region contains a CCAAT-like motif and is important for basal-level expression, and the 3′-half region is important for A23187 and thapsigargin inducibility (19, 22). The same region defined through mutagenesis as important for stress induction also contained the sequence where changes in factor occupancy were observed. Next, we used gel mobility shift assays to determine whether we can detect in HeLa nuclear extracts binding activities which are specific for 3′ half of the core. Using nuclear extracts prepared from A23187-induced cells and the rat *grp* core sequence as the probe, four complexes (I, II, II′, and III) were observed (Fig. 4C). Previously we

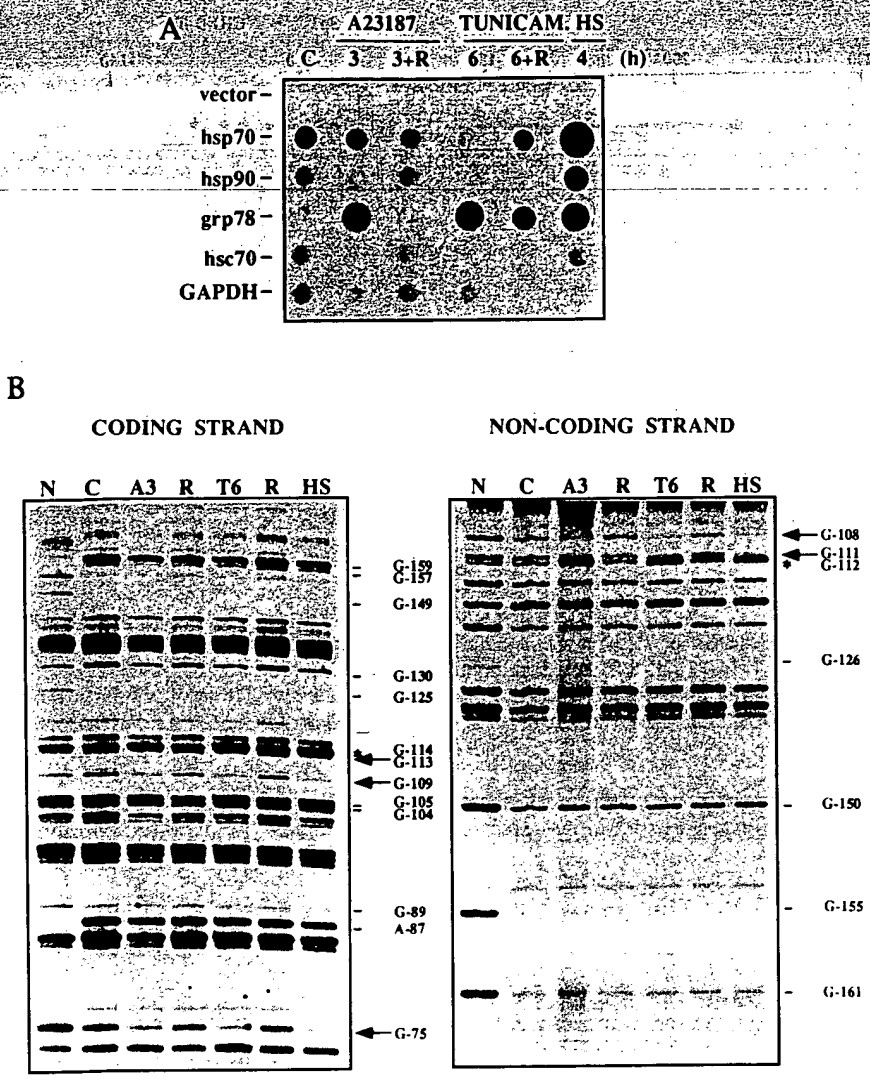


FIG. 2. (A) Transcription activation of the *grp78* gene during A23187 and tunicamycin treatments or heat shock of HeLa cells. Nuclear run-on analyses of *hsp70*, *hsp90α*, *grp78*, *hsc70*, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) transcription were performed at the indicated time after addition of A23187 and tunicamycin or exposure to 42°C (heat shock [HS]). Vector denotes plasmid pGEM2. Lanes R represent cells that after being treated with A23187 and tunicamycin were washed and allowed to recover for 6 h before the nuclei were isolated. (B) In vivo genomic footprint of the core region of the human *grp78* promoter. After DMS treatment in vivo, genomic DNA was isolated from the parallel samples that were used for the nuclear run-on analysis, and methylation patterns of G residues were compared between control cells (C) and cells that were exposed to A23187 (A), tunicamycin (T), or heat shock (HS) treatment. Arrows indicate the G residues that are protected from methylation, and asterisks indicate the G residues that are hypersensitive to methylation.

demonstrated that a Ku autoantigen found in high abundance in HeLa nuclear extracts has a high affinity for the core binding site and is a major component of complex II (22). With modifications in the gel mobility shift assay and an extended electrophoresis running time, we discovered a faster-migrating complex, II′, which can be resolved from the major complex II.

To dissect the binding domains of the four complexes, competition assays were performed with the synthetic oligomers of the rat *grp* core mutated at either the 5′ or 3′ half (Fig. 4B). To control for nonspecific binding, a random synthetic oligomer bearing the mutated sequences of both the 5′ and 3′ halves was used. With the exception of the mutated bases, the sizes and the linker tails of all of the synthetic oligomers are identical. As shown in Fig. 4C, upon competition with 100- and 200-fold molar excesses of the synthetic oligomers described above, complex II′ corresponds to the complex which exhibits high affinity for the *grp* core 3′-half region, as complex II′ was unaffected by the synthetic oligomers with the 3′-half mutation or the random sequence competitors but was efficiently eliminated by the wild-type core or the 5′-half mutation competitors. In contrast, complex I and the majority of complex II appeared to be nonspecific, and complex III could be competed for equally by both the 5′ and 3′ halves of the core.

In vitro core binding activities in control and induced HeLa nuclear extracts. To test whether the in vivo protection pattern changes correlate with changes of in vitro binding for the *grp* core, nuclear extracts were prepared from spinner cultures of exponentially growing HeLa cells maintained in normal culture medium or treated with the ionophore A23187 for 5 h. RNA blot analysis of these cells routinely showed a 10-fold increase in *grp78* mRNA levels in the A23187-treated cells (Fig. 5A),

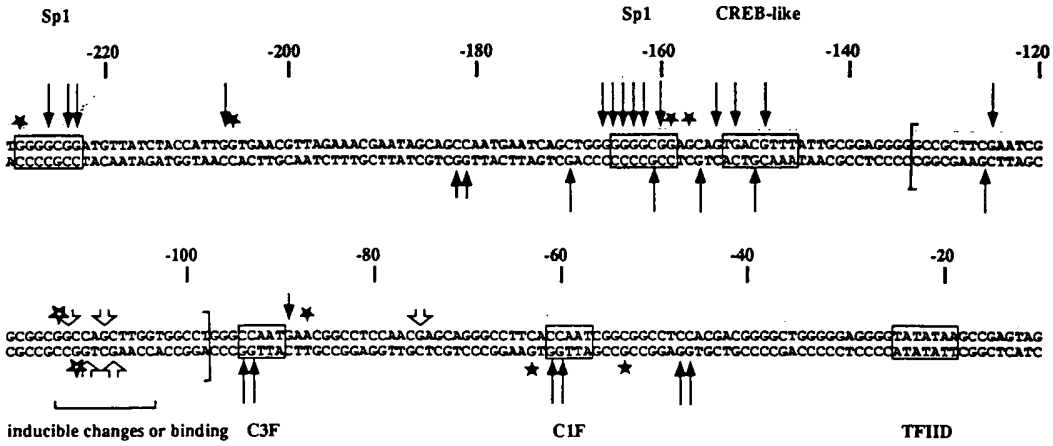


FIG. 3. Schematic presentation of the human *grp78* promoter containing the consensus sequences and the respective transcription factors bound to these sites. Closed arrows and stars denote constitutive protections and hypersensitive sites of the DMS methylation of G residues, respectively, as described in the legend to Fig. 1, and open arrows and stars denote inducible changes in the methylation pattern.

demonstrating that the control cells were not subjected to ER stress during culture or harvest. The rRNA profile was used as a control for even loading of mRNA. With these nuclear extracts and the rat *grp* core sequence as the probe in gel mobility shift assays, the binding activities of the four complexes (I, II, II′, and III) from the control and A23187 extracts were indistinguishable (Fig. 5B).

To characterize further the kinetics of formation of these complexes from control and induced extracts, the rates of association and dissociation of the complexes were measured. We observed that the formation of all four complexes was rapid such that steady-state levels of the binding activities were reached within 5 min of incubation of the nuclear extracts with the core probe (data not shown). In dissociation rate measurements, the extracts were preincubated with the core probe to allow the formation of the complexes prior to the addition of a 50-fold molar excess of the unlabeled core oligomers for various times (5, 15, 30, and 40 min) after the preincubation

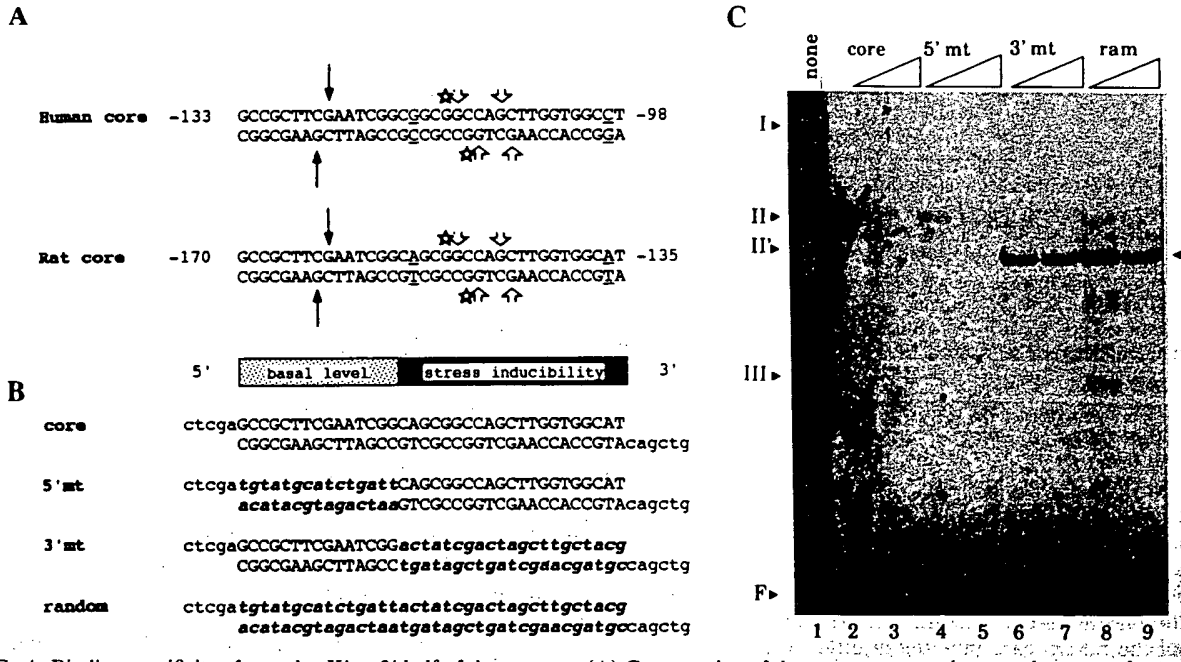


FIG. 4. Binding specificity of complex II′ to 3′ half of the *grp* core. (A) Conservation of the core sequences between human and rat *grp78*. The relative position of the core sequence within each promoter is indicated. The divergent base pairs between the two core sequences are underlined. The bases which were protected in vivo in control cells are denoted by solid arrows; those bases which exhibited changes in site occupancy after stress induction are marked by open arrows (protection) and open stars (hypersensitivity). (B) Sequences of the oligomers used as probes and competitors. The wild-type sequence of the rat *grp78* core is shown in capital letters. Bold italic letters represent the mutated sequences, and lowercase letters indicate the linker sequences. (C) Competition studies using the core oligomer as a probe. Each reaction mixture contained the human nuclear extract mixed with the *grp78* core probe and either no competitors (lane 1) or a 100- to 200-fold molar excess of the respective competitors as indicated (lanes 2 to 9). The autoradiograms are shown. The positions of the complexes and the free probe (F) are indicated.

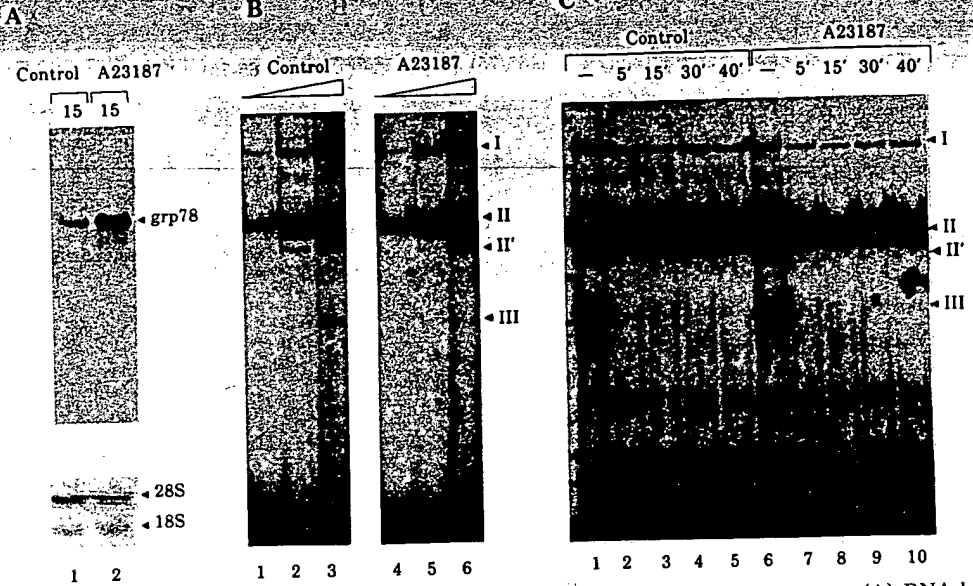


FIG. 5. In vitro core binding activities and kinetics in control and A23187-treated HeLa nuclear extracts. (A) RNA blot analysis of *grp78* transcript levels in control and A23187-treated HeLa cells used for nuclear extract preparation. Fifteen micrograms of each RNA sample was applied and hybridized with the *grp78* cDNA probe. The *grp78* mRNA bands are indicated. (B) Comparison of the core binding activities from control and A23187-induced HeLa nuclear extracts. Gel mobility shift assays were performed with the rat *grp78* core as the probe and 0.1, 0.25, and 0.5 μg of either control or induced HeLa extracts. The autoradiograms are shown. The positions of the four DNA-protein complexes (I, II, II′, and III) are indicated. (C) Comparison of the dissociation rates of the core binding activities. The rat *grp78* core probe was first incubated with either control or induced nuclear extract for 20 min. Except for lanes 1 and 6, a 50-fold molar excess of unlabeled core oligonucleotides was added, and the reaction mixtures were further incubated for the indicated period of time (in minutes) before gel shift analysis.

period. The results of these experiments (Fig. 5C) revealed that complexes I and II were highly stable. Once they were formed, there was no exchange with the unlabeled competitors during the 40-min incubation period. In contrast, complexes II′ and III exhibited a rapid dissociation rate. Within 5 min of incubation with the unlabeled competitor, complex II′ was no longer detected. Incubation of unlabeled competitors at a 100-fold molar excess for as short as 1 min also effectively eliminated complex II′ (data not shown). For both control and induced extracts, the dissociation rates for complexes II and III were extremely rapid, whereas those for complexes I and II were slow. Since complex II′ binds with the highest specificity to the 3′ half of the core, which is functionally most significant because of its contribution to stress induction, we focused on identifying the protein component of complex II′ which binds to this DNA region.

Complex II′ can be reconstituted from renatured protein species of 65 to 75 kDa. Using UV cross-linking techniques, we identified the sizes of several polypeptides (210, 110, 90/92, and 70 kDa) from HeLa nuclear extracts which could be photolabeled within the major complex II (including that of II′), with a 70-kDa protein species being the most predominantly labeled band (22). As the first step to ascertain whether any of the protein species identified above has the ability to bind directly to the *grp* core, HeLa nuclear extracts prepared from A23187-treated cells were size fractionated in preparative SDS-polyacrylamide gels. The gels were cut into 1-cm slices, and the proteins, after electroelution, were allowed to renature. The renatured proteins were then used in gel mobility shift assays using the rat *grp* core as the probe. Examples of the gel mobility shift assays performed with renatured proteins recovered from the molecular size ranges of 115 to 140, 90 to 115, and 65 to 75 kDa are shown in Fig. 6. Using the regular binding buffer which contains Mg^{2+} as the divalent ion, the only

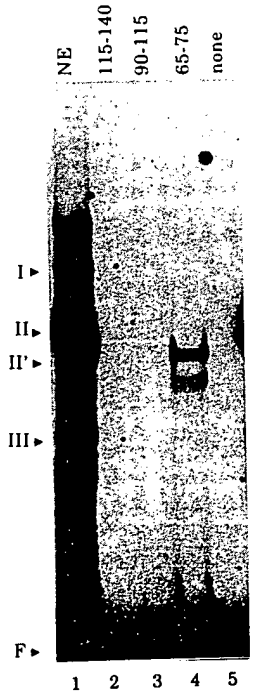


FIG. 6. Renaturation of complex II′ binding activity. Representative renatured fractions from regions of the denaturing SDS-polyacrylamide gel were assayed for core binding activities. Each reaction mixture contained the rat *grp78* core probe and either 0.1 μg of induced nuclear extract (NE; lane 1), 2 μl of renatured fractions (lanes 2 to 4), or probe only (lane 5) and was incubated on ice for 2 h. The positions of the complexes and free probes (F) are indicated. The size range of proteins from each renatured fraction is indicated in kilodaltons at the top.

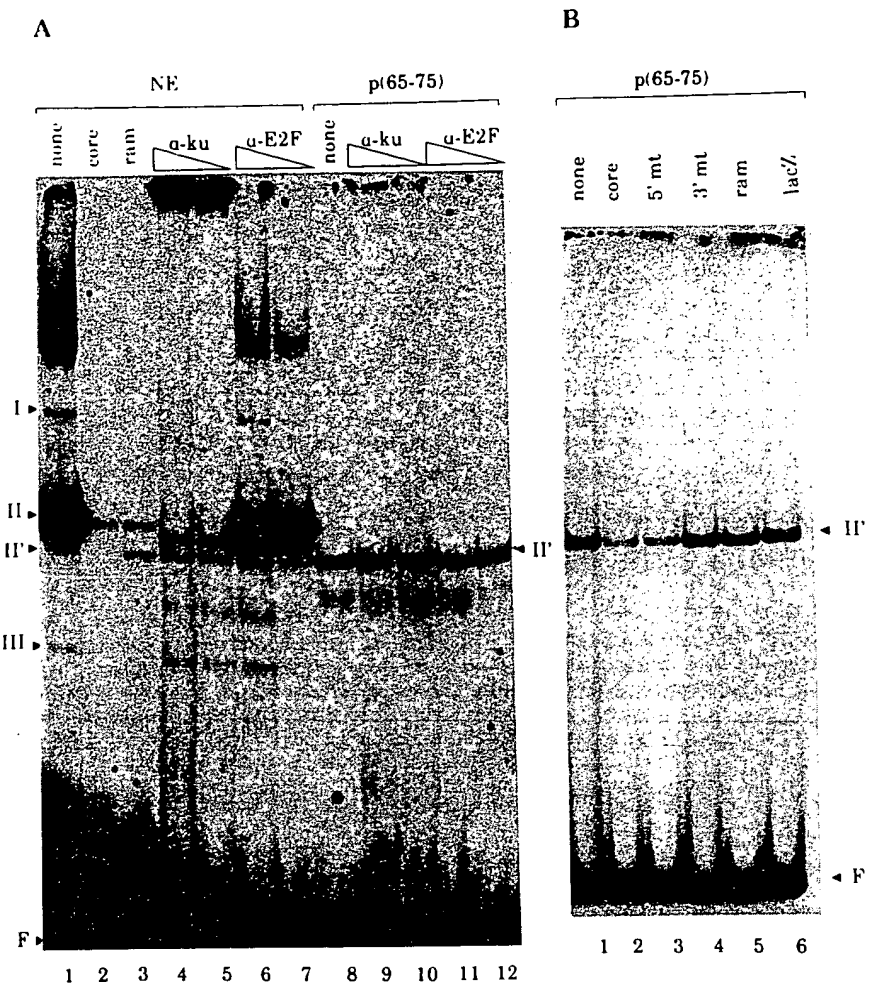


FIG. 7. Specificity of the renatured complex II′ activity from p(65-75). (A) Gel shift analysis shows that p(65-75) is devoid of Ku activity. In lanes 1 to 7, each reaction mixture contained the rat *grp78* core probe, 0.1 μg of induced nuclear extract, and the following: no competitors (lane 1), 100-fold molar excess of unlabeled core oligomers (lane 2), 100-fold molar excess of random competitors (lane 3), 0.1 to 0.05 μl of anti-Ku antiserum (αku; lanes 4 and 5), or 0.1 to 0.05 μl of an unrelated anti-E2F antiserum (α-E2F; lanes 6 and 7). In lanes 8 to 12, each reaction mixture contained the rat *grp78* core probe, 2 μl of the renatured p(65-75) fraction, and the following: no antiserum (lane 8) or 0.1 to 0.05 μl of anti-Ku (lanes 9 and 10) or anti-E2F (lanes 11 and 12) antiserum. The antibodies were preincubated with the nuclear extract for 3 to 5 min before addition of the probe. All reaction mixtures were incubated further on ice for 20 min. (B) Specificity of the renatured p(65-75) binding activity. Competition studies were performed with a fivefold molar excess of competitors depicted in Fig. 4B. The positions of the free probe (F) and complex II′ are marked.

complex which could be reconstituted was complex II′, and that was achieved with renatured proteins in the size range of 65 to 75 kDa. In an attempt to stabilize any DNA-binding factor which may require Zn^{2+} for renaturation and binding, we tested whether addition of Zn^{2+} to the binding buffer could enhance binding to the core from any of the fractions. We observed that the binding activities were only slightly enhanced by the presence of Zn^{2+} (data not shown). These results with the unfractionated HeLa nuclear extract indicate that the protein species between 65 and 75 kDa, referred to below as p(65-75), has the ability to bind to the *grp* core after renaturation. While complex II′ is most readily reconstituted from renatured protein fractions, it is possible that there are other species which also bind to the core probe but were not able to renature.

The renatured protein p(65-75) is distinct from the Ku autoantigen and binds specifically to the *grp* core. Previously, the abundant complex II was shown to contain a Ku autoantigen binding activity and a core-specific binding activity (22). Ku binds to double- and single-stranded DNA ends with high affinity and random sequence specificity (1). The Ku autoantigen consists of a 70-kDa DNA-binding component and a 80-kDa component. Since the protein species that we identified which can reconstitute complex II′ are similar in molecular size to the DNA-binding component of Ku, it is critical to establish that p(65-75) is not Ku or contaminated with it. For this purpose, we used two approaches to test the identity and specificity of p(65-75). The first approach was to test whether p(65-75) can react with the antibody against Ku and, in doing so, affect complex II′ formation with the *grp* core. As shown in Fig. 7A, the abundant complex II formed with the unfractionated HeLa extract was largely supershifted to the gel wells by the anti-Ku antibody, whereas an unrelated anti-E2F antibody had no effect. In contrast, the formation of the reconstituted complex II′ derived from renatured p(65-75) was not affect-

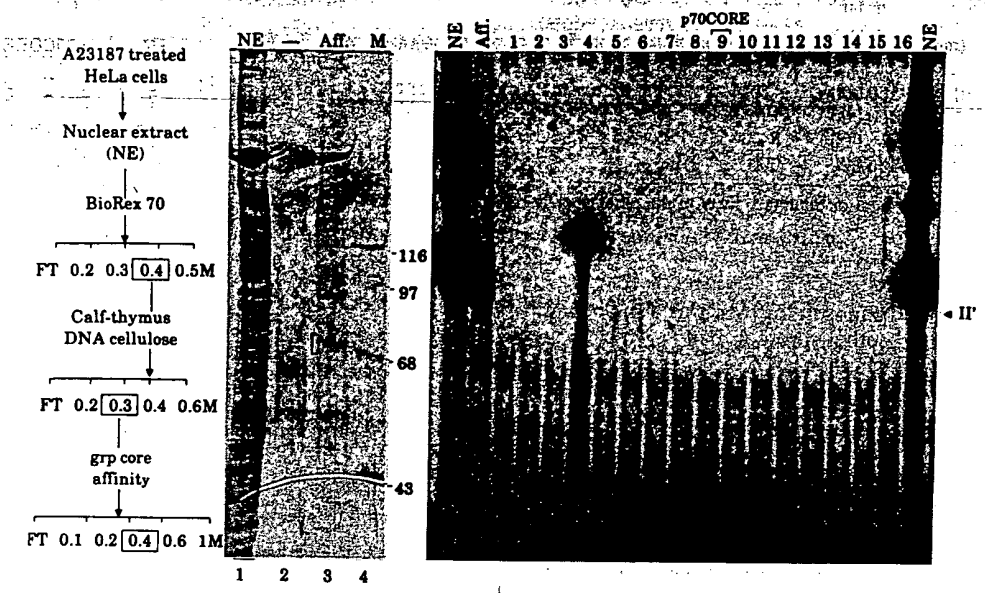


FIG. 8. Chromatographic purification of complex II′ binding activity and identification of p70CORE as the core-binding polypeptide. (A) Column purification scheme. The activity of complex II′ was monitored by gel shift assays. The salt concentrations at which II′ activity eluted from various columns are boxed. FT, flowthrough. (B) Silver stain profile of the core affinity-purified II′ binding activity. Lane 1, unfractionated nuclear extract (NE); lane 2, buffer alone; lane 3, 0.4 M eluate from affinity-purified fraction (Aff.); lane 4, molecular size markers (M). The bracketed cluster of bands around 70 kDa is referred to as p70CORE. The sizes of the markers are indicated in kilodaltons. (C) Reconstitution of complex II′ binding activity from renatured p70CORE of the affinity-purified fraction. Each reaction mixture contained the rat *grp78* core probe and either induced nuclear extract (NE), 2 μl of affinity-purified fraction (Aff.), or 20 μl of renatured fractions (1 through 16; the numbers represent the gel slices in the direction of decreasing molecular weight). The reactions were incubated on ice for 2 h before electrophoresis at 4°C. After electrophoresis, the dried gel was exposed for 7 days at −70°C. The dark spot in the middle of lane 4 is an artifact from electrophoresis.

by preincubation with either the anti-Ku antigen or the anti-E2F antibody. These results demonstrate that p(65-75) is not Ku and is not contaminated with it.

Next we determined whether p(65-75), after isolation from a denaturing gel and subsequent renaturation, retained its specificity for the 3′ half of the *grp* core. For this purpose, competition assays with various synthetic oligomers (Fig. 4B) were performed with the reconstituted complex II′ (Fig. 7B). We observed, as in the case of the complex II′ formed from the unfractionated extract (Fig. 4C and 7A, lanes 1 to 3), that the reconstituted complex II′ is highly specific for the 3′ half of the core sequence. Only the wild-type core oligomer and the 5′ mutated oligomer, containing the wild-type 3′-half core sequence, were able to abolish complex II′. The addition of the 3′ mutated core oligomer and two other heterologous competitors had no effect. The specificity of the renatured complex was retained when up to 200-fold molar excesses of competitors were used (data not shown).

Purification of complex II′ binding activity from A23187-induced HeLa extracts. With the establishment of a gel mobility shift assay which allows us to monitor the complex II′ binding activity, the induced HeLa nuclear extract was subjected to a series of chromatographic steps for the purpose of purifying p(65-75), which binds to the 3′ half of the *grp* core with high specificity and affinity, using the purification scheme shown in Fig. 8A.

About 50 mg of HeLa nuclear extract was first applied onto a BioRex 70 column, which is a high-capacity, weak cationic exchange column. Upon step elution with increasing concentrations of KCl, the complex II′ binding activity, as detected by gel mobility shift assays, was enriched in the 0.4 M fractions. The active fractions were then applied onto a double-stranded calf thymus DNA-cellulose column to resolve the DNA-binding proteins on the basis of their relative affinities for general DNA. This column was useful in separating the complex II binding activity from the complex II′ binding activity, which eluted primarily in the 0.3 M fraction. Complex II (mostly Ku) was enriched in later elution fractions (data not shown). Finally, the complex II′ active fractions were applied onto a Sepharose 4B column coupled with the concatenated core oligomers. The bound proteins were eluted with a stepwise KCl gradient, and complex II′ was recovered in the 0.4 M fraction. As shown in Fig. 8C, the affinity-purified fraction in gel mobility shift assays formed only complex II′ with the *grp* core probe. Thus, the multiple chromatographic fractionation steps were successful in segregating the components comprising complexes II and II′. A comparison of the silver stain patterns of the unfractionated extract and the affinity-purified fraction enriched for complex II′ is shown in Fig. 8B. For the affinity-purified fraction, there were several groups of protein bands, including those with sizes in the ranges of 120, 97, and 70 kDa, which were enriched. A lane consisting of the buffer alone was included to account for the cluster of nonspecific bands below the 68-kDa size marker. Based on their relative specific DNA binding activities, the enrichment for complex II′ binding activity at this final stage of chromatographic purification is about 10,000-fold (Table 1).

Identification of p70CORE as the DNA-binding component of complex II′. Using unfractionated HeLa nuclear extract, we demonstrated above that the protein component of complex II′ has a molecular size of about 70 kDa. Upon chromatographic purification, a set of proteins of about 70 kDa (referred

TABLE 1. Purification of p70CORE from HeLa nuclear extract

Fraction	Protein (mg)	Total activity (U)[a]	Sp act (U/μg)	Purification (fold)	Recovery (%)
Nuclear extract	50	500,000	10		
BioRex 70	2.1	400,000	190	19	80
Calf thymus DNA	0.34	288,000	847	85	58
2× core affinity	0.002	186,000	93,000	9,300	37

[a] One unit is defined as the amount of binding activity of 0.1 μg of nuclear extract under normal binding conditions.

to as p70CORE; bracketed bands in Fig. 8B) was also detected by silver staining, among other protein bands which were copurified. To determine whether p70CORE is the purified form of p(65-70) and whether other core-binding proteins were also purified, the affinity-purified fraction containing the complex II′ binding activity was applied onto a preparative SDS-polyacrylamide gel to fractionate the proteins. Gel slices of 1 cm each were cut along the whole length of the gel. After elution, the proteins were allowed to renature and used in gel mobility shift assays, using the rat *grp* core as the probe. The results are shown in Fig. 8C. Of all of the gel slices tested, only slice 9, containing p70CORE, was able to form a complex. The complex is II′, as judged by its electrophoretic mobility.

To confirm that the affinity-purified fraction and p70CORE retained the ability to bind to the 3′ half of the *grp* core, competition assays with wild-type and mutated core oligomers were performed as described for Fig. 4C. As shown in Fig. 9A, a side-by-side comparison of the electrophoretic mobility of complex II′ formed by the affinity-purified fraction and that of the unfractionated nuclear extract confirmed that they were identical. Complex II′ as formed by the affinity-purified fraction was specifically inhibited by the *grp* core oligomer but not by the random oligomer competitor. Whereas p70CORE after its isolation and renaturation exhibits a substantial loss of protein mass and binding activity (only ~1% recovery of binding activity), a side-by-side comparison of complex II′ derived from the affinity-purified fraction and p70CORE showed that they migrated with identical electrophoretic mobilities (Fig. 9B). Further, the specificity of p70CORE for the 3′ half of the *grp* core was retained, as demonstrated by competitions with the wild-type and 5′ mutated oligomers but not the 3′ mutated or random oligomers at a 100-fold molar excess (Fig. 9B).

To determine whether the affinity-purified fraction binds to the same bases which exhibit in vivo stress-inducible changes in methylation pattern, competition analyses with additional synthetic oligonucleotides were performed. The mutations, M1 through M4, each contained 4-bp substitutions within the 3′ half region of the *grp* core (Fig. 10A). The complex II′ formed by the affinity-purified fraction can be competed for only by the wild-type core and M4 oligomers, which mutated 4 bp outside the in vivo DMS footprint region (Fig. 10B). The oligomers M1 and M2, which created base mutations within the DMS-protected region, were unable to compete efficiently even at a 100-fold molar excess. The inability of M3 to compete suggests that bases undetectable by DMS protection in vivo in the immediate flanking region may also be involved in protein-DNA interaction. In comparison, the competitor with the least potency is 3′mt, which mutated the entire 3′ half of the *grp* core. These combined results demonstrate that the affinity-purified fraction binds to the same sites which exhibit stress-inducible changes in DMS protection and that p70CORE as purified from the HeLa nuclear extract is the DNA-binding component of complex II′ which binds specifically to the region of the *grp* core required for stress induction.

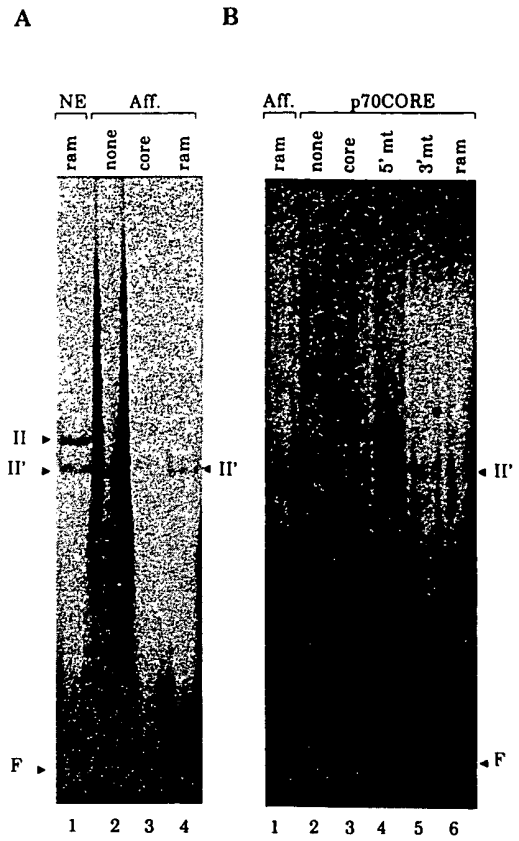


FIG. 9. Specificity of the p70CORE binding activity. (A) Specificity of the affinity-purified fraction as determined from competition studies. Lane 1 contains the rat *grp78* core probe mixed with induced nuclear extract (NE) and a 100-fold molar excess of random competitors. Lanes 2 to 4 contain the rat *grp78* probe and 2 μl each of the affinity-purified II′ binding activity (Aff.) and either no competitors (lane 2) or a 100-fold molar excess of unlabeled core (lane 3) or random (lane 4) competitors. (B) Specificity of the renatured p70CORE binding activity. Lane 1 is a longer exposure of lane 4 in panel A. Lanes 2 to 6 contain the rat *grp78* core probe, 20 μl each of renatured p70CORE and a 100-fold molar excess of the various competitors as indicated at the top. After electrophoresis, the dried gel was exposed for 14 days at −70°C. The smearing effect observed in lane 2 of both panels A and B is due to the 12.5% glycerol in the final 40-μl reaction volume. F, free probe.

DISCUSSION

The regulation of GRP78 in response to stress conditions in the ER represents a unique model for signal transduction. In the case of the mammalian *grp* genes, a depletion of the calcium store from the ER or the accumulation of malfolded or underglycosylated proteins apparently generates a signal that is transmitted to the nucleus and leads to enhanced transcription of the *grp* genes. This transcriptional regulation is mediated primarily by upstream promoter sequences (19, 32). While the sequence organization of the rat *grp78* promoter is complex and contains functionally redundant elements (47), the sequence spanning −170 to −134, termed the *grp* core, is strikingly conserved in *grp78* and *grp94* promoters in speci

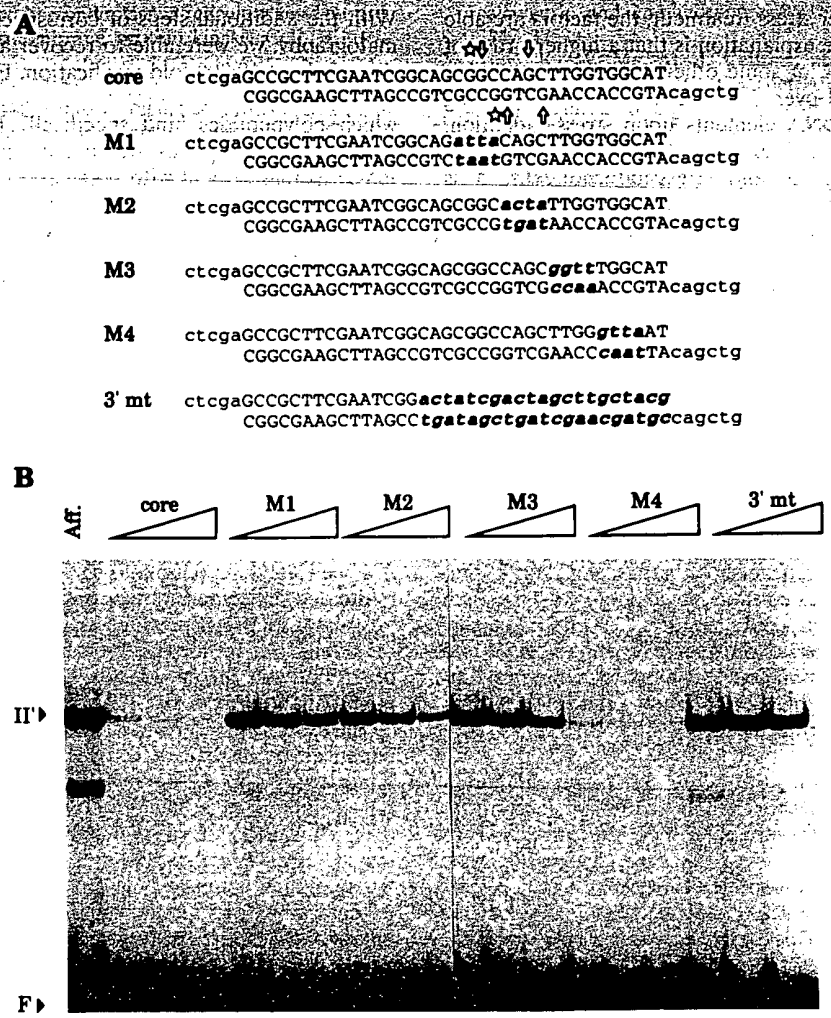


FIG. 10. Delimitations of the complex II′ binding site with the affinity-purified fraction. (A) Sequences of the oligomers used as probes and competitors. The wild-type sequence of the rat *grp78* core is shown in capital letters, with the bases which exhibited changes in site occupancy after stress induction marked by arrows (protection) and stars (hypersensitivity). Bold italic letters indicate the mutated sequences, and lowercase letters denote the linker sequence. (B) Competition of complex II′. Each lane contained 2 μl of the affinity-purified fraction (Aff.) with either no competitor (first lane) or increasing molar excesses (25-, 50-, and 100-fold) of the competitors indicated at the top. The autoradiograms are shown. The positions of complex II′ and the free probe (F) are indicated.

ranging from yeasts to humans (5, 32). This high degree of conservation is consistent with its functional significance in ER stress induction, as revealed by promoter mapping studies (22) and its ability to confer ER stress inducibility to heterologous promoters in both the mammalian and yeast systems (19, 28). Moreover, binding element titration in vivo reveals that the core element is essential for stress induction of the *grp* genes and for cell survival (5, 20), and the subdomain responsible for calcium stress inducibility was mapped to the 3′ half of the core region (22).

To understand how the core element mediates stress induction of *grp78* transcription, we have previously performed gel mobility shift assays with a 291-bp *grp* promoter fragment (7) to determine if differences in factor binding could be detected in induced cell extracts compared with control cell extracts. In those in vitro assays, the binding activities were similar. Further, in yeast cells, the homologous *grp* core binding activity was also reported to be constitutive, and no differences between control and stressed cells were found (28). Using in vivo footprinting techniques to investigate protein-DNA interactions of the human *grp78* promoter, we report here that specific changes in the methylation pattern such as protection or hypersensitivity of specific nucleotides were detected on the coding and noncoding strands within the 3′ half of the *grp* core when the cells were stressed. Further, the changes in factor occupancy directly correlate with transcriptional induction of the endogenous *grp78* gene. Nonetheless, the lack of in vivo protection in nonstressed cells was apparently not due to the absence of factors that are capable of binding to the highly conserved *grp* core, since extracts from induced and control cells bind with similar specificities and kinetics to the rat *grp* core, as analyzed in vitro by gel mobility shift assay. The lack of in vivo footprinting despite the presence of binding factors has been reported for the promoters of the tyrosine aminotransferase (4), major histocompatibility complex class II (14), and β_2-microglobulin genes (25). Thus, the *grp78* gene provides an additional example wherein the presence of element-specific factors in control cells is not sufficient for producing in vivo

occupancy, whereas after stress treatment, the factors are able to access the DNA. One explanation is that a higher order of chromatin configuration, or some other inducible interacting factor(s), exerts control over the accessibility of the binding factor to the specific DNA elements upon stress induction. Since similar binding activities were detected in vitro from nuclear extracts of control and stress-induced cells, it is unlikely that the accessibility is controlled by nuclear transport of a cytoplasmic factor. The combined results of the in vitro and in vivo analyses underscore the importance of examining the DNA-protein interactions in vivo in the context of the chromatin, where subtle differences in factor binding would otherwise be undetectable in vitro.

Further, we noted that despite the lack of a heat shock element, identical protections and hypersensitivities of guanine residues were seen in the human *grp78* promoter from heat-shocked cells and those treated with A23187 or tunicamycin. The *grp78* gene (*KAR2*) in yeast cells is heat shock inducible; however, this response is mediated by a heat shock element within the yeast *KAR2* promoter (28). One explanation is that the kinetics of heat induction of *grp78* is distinct from the kinetics of other classical heat shock genes, e.g., *hsp70* and *hsp90*, such that it requires continuous heat shock (4 h). Such prolonged heat treatment could result in a physiologically stressful condition similar to those elicited by A23187 and tunicamycin. These results indicate that several different stimuli mediate their effects through common inducible protein-DNA interactions occurring at the 3′ half of the *grp* core sequence.

Given the importance of the *grp* core, we sought to identify the binding activity specific for 3′ half of the core region. By using half-site mutants, we were able to define complex II′ as the activity specific for 3′ half of the core. Through kinetic studies, we determined that complex II′ binding to the 3′ half of the core is extremely rapid. At the same time, it dissociates from the core rapidly. This quick exchange rate does not compromise the stability of the complex II′ over time since 5′ mutant and random competitors do not compete even at a high molar excess. This unique property of complex II′ may be favorable for the regulatory role that it portends. It may be one mechanism through which a gene can be quickly attenuated through a change in the accessibility of DNA.

From fractionation of induced nuclear extract, we found that the complex II′ activity may be reconstituted from renatured fractions of polypeptides in the range of 65 to 75 kDa. Since the Ku autoantigen has been previously shown to bind the *grp* core (22), it was necessary to determine whether the renatured activity was specific to the *grp78* core and whether the renatured activity was from the abundant Ku autoantigen. The Ku autoantigen consists of a 70-kDa DNA-binding component and a 80-kDa component. While its binding specificity appears to be nonspecific, Ku was recently shown to be a substrate for a DNA-activated protein kinase (18). Interestingly, the transcription of *grp78* was suppressed by protein kinase inhibitors (30, 33). Whether the binding of Ku to the *grp* core exists in vivo and whether it plays any specific role in activating the core complexes remain to be determined. Nonetheless, we demonstrate here that complex II′ is distinct from the Ku autoantigen. Complex II′ from renatured fractions exhibits the same electrophoretic mobility as that from the unfractionated nuclear extract, suggesting that it is unlikely to consist of heterologous subunits outside the size range of 65 to 75 kDa.

Since complex II′ is highly specific for the 3′ half of the core, we have focused on its purification. Through the use of the BioRex 70 and calf thymus DNA columns (40), we were able to purify the complex II′ activity away from complex II activity. With the additional steps of core-specific DNA affinity chromatography, we were able to recover an affinity-purified fraction of about 10,000-fold purification. From silver stains, these fractions contained several major polypeptides. To determine which polypeptides bind specifically to the *grp78* core, we renatured polypeptides of specific size ranges after SDS-PAGE. In agreement with results of a similar study using the crude nuclear extract, we observed that proteins in the size range of 65 to 75 kDa give renatured binding activity. In this range, the enriched protein bands were a cluster of several polypeptides with molecular sizes of about 70 kDa (p70CORE). By itself, p70CORE reconstitutes the complex II′ binding activity and exhibits binding specificity. Through in vitro competition with the affinity-purified fraction, complex II′ interacts with the same bases for which inducible changes are observed in vivo. The identification of the p70CORE which binds to the region of the core responsible for inducible expression is an important first step toward characterization of the protein involved in and understanding the mechanism of the induction of *grp78*.

Recently, a transmembrane serine/threonine kinase, IREp, has been shown in yeast cells to be essential in the induction by malfolded protein accumulation in the ER (6, 27). It has been postulated that IREp, which spans the ER membrane, transmits a signal from the ER to the nucleus. However, the ligands of IREp, as well as its substrates, are unknown. For the mammalian *grp* system, there are reports that its induction is sensitive to W7, a serine/threonine kinase inhibitor (33), and enhanced by okadaic acid, which is a serine/threonine phosphatase inhibitor (31). However, genistein, a protein tyrosine kinase inhibitor, is also reported to inhibit *grp78* induction (31). Thus, the pathways which regulate mammalian *grp* genes may be more complex than those in yeast cells. Assuming that a mammalian homolog of IREp does exist, it will be important to determine the relationship between p70CORE and this kinase and identify the intermediary signal transduction steps between the ER and the nucleus.

ACKNOWLEDGMENTS

We thank Deborah Johnson for helpful discussion and Eva Liu for establishing the parameters for the chromatographic purification of core-binding factors. We also thank D. M. Livingston for the E2F antiserum and W. H. Reeves for the Ku antiserum.

This research was supported in part by Public Health Service grants R37 CA27607 (to A.S.L.) and GM38109 (to R.I.M.). L.S. was supported by fellowships from the Academy of Finland and the Fogarty International Foundation.

REFERENCES

1. **Allaway, G. P., A. A. Vivino, L. D. Kohn, A. L. Notkins, and B. S. Prabhakar.** 1990. Characterization of the 70 kDa component of the human Ku autoantigen expressed in insect cell nuclei using a recombinant baculovirus vector. Biochem. Biophys. Res. Commun. **168**:747–755.
2. **Baeuerle, P. A., and D. Baltimore.** 1988. Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-κB transcription factor. Cell **53**:211–217.
3. **Banerji, S. S., N. G. Theodorakis, and R. I. Morimoto.** 1984. Heat shock-induced translational control of HSP70 and globin synthesis in chicken reticulocytes. Mol. Cell. Biol. **4**:2437–2448.
4. **Becker, P. B., S. Ruppert, and G. Schuetz.** 1987. Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. Cell **51**:435–443.
5. **Chang, S. C., A. E. Erwin, and A. S. Lee.** 1989. Glucose-regulated protein (GRP94 and GRP78) genes share common regulatory domains and are coordinately regulated by common *trans*-acting factors. Mol. Cell. Biol. **9**:2153–2162.
6. **Cox, J. S., C. E. Shamu, and P. Walter.** 1993. Transcriptional

induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. Cell **73**:1197–1206.

7. **Day, A. R., and A. S. Lee.** 1989. Transcriptional regulation of the gene encoding the 78-kD glucose-regulated protein GRP78 in mouse Sertoli cells: binding of specific factor(s) to the GRP78 promoter. DNA **8**:301–310.
8. **Drummond, I. A. S., A. S. Lee, E. J. Resendez, and R. A. Steinhardt.** 1987. Depletion of intracellular calcium stores by calcium ionophore A23187 induces the genes for glucose-regulated proteins in hamster fibroblasts. J. Biol. Chem. **262**:12801–12805.
9. **Fort, P., L. Marty, M. Piechaczyk, S. S., C. Dani, P. Jeanteur, and J. M. Blanchard.** 1985. Various rat adult tissues express only one major mRNA species from the glyceraldehyde-3-phosphate-dehydrogenase multigenic family. Nucleic Acids Res. **13**:1431–1442.
10. **Gething, M. J., and J. Sambrook.** 1992. Protein folding in the cell. Nature (London) **355**:33–45.
11. **Gomer, C. J., A. Ferrario, N. Rucker, S. Wong, and A. S. Lee.** 1991. Glucose regulated protein induction and cellular resistance to oxidative stress mediated by porphyrin photosensitization. Cancer Res. **51**:6574–6579.
12. **Hickey, E., S. E. Brandon, G. Smale, D. Lloyd, and L. A. Weber.** 1989. Sequence and regulation of a gene encoding a human 89-kilodalton heat shock protein. Mol. Cell. Biol. **9**:2615–2626.
13. **Kadonaga, J. T., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA **83**:5889–5893.
14. **Kara, C. J., and L. H. Glimcher.** 1991. *In vivo* footprinting of MHC class II genes: bare promoters in the bare lymphocyte syndrome. Science **252**:709–712.
15. **Kozutsumi, Y., M. Segal, K. Normington, M. J. Gething, and J. Sambrook.** 1988. The presence of malfolded proteins in the endoplasmic reticulum signals the induction of glucose-regulated proteins. Nature (London) **332**:462–464.
16. **Lee, A. S.** 1987. Coordinated regulation of a set of genes by glucose and calcium ionophores in mammalian cells. Trends Biochem. Sci. **12**:20–23.
17. **Lee, A. S.** 1992. Mammalian stress response: induction of the glucose-regulated protein family. Curr. Opin. Cell Biol. **4**:267–273.
18. **Lees-Miller, S. P., Y. R. Chen, and C. W. Anderson.** 1990. Human cells contain a DNA-activated protein kinase that phosphorylates simian virus 40 T antigen, mouse p53, and the human Ku autoantigen. Mol. Cell. Biol. **10**:6472–6481.
19. **Li, W. W., S. Alexander, X. Cao, and A. S. Lee.** 1993. Transactivation of the grp78 promoter by Ca^{2+} depletion. A comparative analysis with A23187 and the endoplasmic reticulum Ca^{2+}-ATPase inhibitor thapsigargin. J. Biol. Chem. **268**:12003–12009.
20. **Li, X. A., and A. S. Lee.** 1991. Competitive inhibition of a set of endoplasmic reticulum protein genes (GRP78, GRP94, and ERp72) retards cell growth and lowers viability after ionophore treatment. Mol. Cell. Biol. **11**:3446–3453.
21. **Lin, A. Y., S. C. Chang, and A. S. Lee.** 1986. A calcium ionophore-inducible cellular promoter is highly active and has enhancer-like properties. Mol. Cell. Biol. **6**:1235–1243.
22. **Liu, E. S., and A. S. Lee.** 1991. Common sets of nuclear factors binding to the conserved promoter sequence motif of two coordinately regulated ER protein genes, *GRP78* and *GRP94*. Nucleic Acid Res. **19**:5425–5431.
23. **Liu, E. S., J. H. Ou, and A. S. Lee.** 1992. Brefeldin A as a regulator of grp78 gene expression in mammalian cells. J. Biol. Chem. **267**:7128–7133.
24. **Lodish, H. F., N. Kong, and L. Wikstroem.** 1992. Calcium is required for folding of newly made subunits of the asialoglycoprotein receptor within the endoplasmic reticulum. J. Biol. Chem. **267**:12753–12760.
25. **Lonergan, M., A. Dey, K. G. Becker, P. D. Drew, and K. Ozato.** 1993. A regulatory element in the β2-microglobulin promoter identified by in vivo footprinting. Mol. Cell. Biol. **13**:6629–6639.
26. **McCaffrey, P. G., B. A. Perrino, T. R. Soderling, and A. Rao.** 1993. NF-AT_p, a T lymphocyte DNA-binding protein that is a target for calcineurin and immunosuppressive drugs. J. Biol. Chem. **268**:3747–3752.
27. **Mori, K., W. Ma, M. J. Gething, and J. Sambrook.** 1993. A transmembrane protein with a cdc^{2+}/CDC28-related kinase activity is required for signaling from the ER to the nucleus. Cell **74**:743–756.
28. **Mori, K., A. Sant, K. Kohno, K. Normington, M. J. Gething, and J. F. Sambrook.** 1992. A 22 bp cis-acting element is necessary and sufficient for the induction of the yeast KAR2 (BiP) gene by unfolded proteins. EMBO J. **11**:2583–2593.
29. **Mueller, P. R., and B. Wold.** 1989. *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. Science **246**:780–786.
30. **Price, B. D., and S. K. Calderwood.** 1992. Gadd45 and Gadd153 messenger RNA levels are increased during hypoxia and after exposure of cells to agents which elevate the levels of the glucose-regulated proteins. Cancer Res. **52**:3814–3817.
31. **Price, B. D., R. L. A. Mannheim, and S. K. Calderwood.** 1992. Brefeldin A, thapsigargin, and AIF4- stimulate the accumulation of GRP78 mRNA in a cycloheximide dependent manner, whilst induction by hypoxia is independent of protein synthesis. J. Cell. Physiol. **152**:545–552.
32. **Resendez, E. J., J. W. Attenello, A. Brafsky, C. S. Chang, and A. S. Lee.** 1985. Calcium ionophore A23187 induces expression of glucose-regulated genes and their heterologous fusion genes. Mol. Cell. Biol. **5**:1212–1219.
33. **Resendez, E. J., J. Ting, K. S. Kim, S. K. Wooden, and A. S. Lee.** 1986. Calcium ionophore A23187 as a regulator of gene expression in mammalian cells. J. Cell Biol. **103**:2145–2152.
34. **Resendez, E. J., S. K. Wooden, and A. S. Lee.** 1988. Identification of highly conserved regulatory domains and protein-binding sites in the promoters of the rat and human genes encoding the stress-inducible 78-kilodalton glucose-regulated protein. Mol. Cell. Biol. **8**:4579–4584.
35. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
36. **Sambrook, J. F.** 1990. The involvement of calcium in transport of secretory proteins from the endoplasmic reticulum. Cell **61**:197–199.
37. **Shapiro, D. J., P. A. Sharp, W. W. Wahli, and M. J. Keller.** 1988. A high-efficiency HeLa cell nuclear transcription extract. DNA **7**:47–55.
38. **Sistonen, L., K. D. Sarge, B. Phillips, K. Abravaya, and R. I. Morimoto.** 1992. Activation of heat shock factor 2 during hemin-induced differentiation of human erythroleukemia cells. Mol. Cell. Biol. **12**:4104–4111.
39. **Ting, J., and A. S. Lee.** 1988. Human gene encoding the 78,000-dalton glucose-regulated protein and its pseudogene: structure, conservation, and regulation. DNA **7**:275–286.
40. **Treisman, R.** 1987. Identification and purification of a polypeptide that binds to the c-fos serum response element. EMBO J. **6**:2711–2717.
41. **Villa, A., P. Podini, D. O. Clegg, T. Pozzan, and J. Meldolesi.** 1991. Intracellular Ca^{2+} stores in chicken Purkinje neurons: differential distribution of the low affinity-high capacity Ca^{2+} binding protein, calsequestrin, of Ca^{2+} ATPase and of the ER lumenal protein, BiP. J. Cell Biol. **113**:779–791.
42. **Watowich, S. S.** 1990. Thesis. Northwestern University, Evanston, Ill.
43. **Watowich, S. S., and R. I. Morimoto.** 1988. Complex regulation of heat shock- and glucose-responsive genes in human cells. Mol. Cell. Biol. **8**:393–405.
44. **Wong, W. L., M. A. Brostrom, G. Kuznetsov, Y. D. Gmitter, and C. O. Brostrom.** 1993. Inhibition of protein synthesis and early protein processing by thapsigargin in cultured cells. Biochem. J. **289**:71–79.
45. **Wooden, S. K., R. P. Kapur, and A. S. Lee.** 1988. The organization of the rat GRP78 gene and A23187-induced expression of fusion gene products targeted intracellularly. Exp. Cell Res. **178**:84–92.
46. **Wooden, S. K., and A. S. Lee.** 1992. Comparison of the genomic organizations of the rat grp78 and hsc73 gene and their evolution-

ary implications. DNA Sequence **3**:41–48.

47. **Wooden, S. K., L. J. Li, D. Navarro, I. Qadri, L. Pereira, and A. S. Lee.** 1991. Transactivation of the grp78 promoter by malfolded proteins, glycosylation block, and calcium ionophore is mediated through a proximal region containing a CCAAT motif which interacts with CTF/NF-I. Mol. Cell. Biol. **11**:5612–5623.

48. **Wu, B., C. Hunt, and R. I. Morimoto.** 1985. Structure and expression of the human gene encoding major heat shock protein HSP70. Mol. Cell. Biol. **5**:330–341.

49. **Wu, C., S. Wilson, B. Walker, I. Dawid, T. Paisley, V. Zimarino, and H. Ueda.** 1987. Purification and properties of *Drosophila* heat shock activator protein. Science **238**:1247–1253.